





```

670 ThrSerAsnMetHisGlySerLeuLeuGlnGlyHisArgIleLeuH1 686
      |||
2433 .....GGAAGCATGCTCCAG.....ATTCATC 2455
686 slyLeuLeuGlnAsnGlyAsnSerProAlaIleAlaIleThr. 702
      |||
2456 TACTATAGCCCAAAACAGACGCCAGAGTGTGACGACACTTCATC 2505
703 .....AlaGlnAlaThrGlyLysAsp 709
2506 AGCAGACAGCATAGATTTTCTCAAGTCCACAACTTGACCAAAATA 2555
710 ThrSerSerIleThrSerGlyAsp..... 718
      |||
2556 GTGACGCTCTCTGCTGCTGTGGGCGACATGCTAGTACTATGCT 2605
719 .....GlyAsnValVal.....LysGlnGlu 726
      |||
2606 TATGGGCGACAGTGTGACGCTCATCTCTTGTGACACAAACAGCAC 2655
726 LLeuSerProLysLysGlyAsnAsnAlaLeuLeuArgTyrLeuLeu 742
2656 AG..... 2657
743 AspArgAspAspProSerAspAlaLeuSerLysGluLeuGlnProGlnVal 759
      |||
2658 .....TCACAGACATTTCTAGTACACGACGACGACGACGACGA 2692
759 LglGlyValAspAsnLysMetSerGlnCysThrSerSerThrIlePro 776
      |||
2693 GCAG.....AGCTCCAGAGGACGACGCTCAGTTCAGTTCACACAA 2736
776 eSerSerGlnGlyLysAspProLysLysThrGluThrSerGluGlu 792
      |||
2737 CTCAGGCTCAGCTGACCCAGCACA..... 2759
793 GlySerGlyAspLeuAspAsnLeuAspAlaIleLeuGlyAspLeuThrSe 809
2759 ..... 2759
809 rSerAspPheTyrAsnAsnSerIleSerSerAsnGlySerHisLeuGly 826
2759 ..... 2759
826 hIryGlnGlnValPheGlnGlyThrAsnSerLeu...GlyLeuLysSer 841
      |||
2760 ..CCGCACACATTTTACAGACTTCTAGTGTCTCCATGGGAATCCCTCA 2807
842 SerGlnSerValGlnSerIleArgProPro.TyrAsnArgAlaValSerI 858
2808 ACTCAACTCATCTCTGCTGCTGATTCCTCTACAAACAGACACTTCCC 2857
858 eu.....AspSerProValSerValGlySerSerProPro 869
      |||
2858 TCAGTCACATCACCAGACATCATGCTCTCAGCAACAGCAGCACTCAGCC 2907
870 ValLysAsnIleSerAlaPheProMetLeuProLysGlnProMetLeuG1 886
      |||
2908 GGCACAGGAGTGCAGCTTCCCGACCTTCCAGG..... 2943
886 yGlyAsnProArgMetLeuAspSerGlnGluAsnTyrGlySerSerMetG 903
2944 .....TTCACACACATGACACACAGTCTCTCTCTT. 2976
903 LysGlyProAsnArgAsnValThrValThrGlnThrProSerSerGlyAsp 919
2976 ..... 2976
920 TyrGlyLeuProAsnSerLysAlaGlyArgMetGluProMetAsnSerAs 936
      |||
2977 .....GACATCAAGGAGGAAAGGATGCGCCCATTAAGAT.. 3012

```

```

936 nSerMetGlyArgProGlyLysPtyrAsnThrSerLeuProArgProA 953
3013 .....TACTCAGATGACTGAGGAAA..... 3033
953 lAleuGlyGlySerIleProThrLeu.....ProLeuArgSerAsn 966
      |||
3034 .....GGAGGGAAGATTCACAGCTTTCATGACATGACAGATTCATG 3078
967 SerIleProGly.....AlaArgProValLeuGlnGlnGlnGln 980
      |||
3079 CTAGTTCCTGGAATGATTGCGACAGAAAATCCTGCTAGTCTACAGAT 3128
980 nMetLeuGlnMetArgProGlyLysIleIleProMetGlyMetGlyAlaAsn 997
      |||
3129 GTACATTTAATACCAACCCAGCA.....GGAGGTGATC 3160
997 rOTyrGlyGlnAlaAlaAlaSerAsnGlnLeuGlySerTrpPro..... 1011
      |||
3161 ATAGGGGACAGCCAGCTTCGACAGTGTTTAGTGCCTGATATTTT 3210
1012 AspGlyMetLeuSerMetGluGlnValSerHisGlyThrGlnAsnArgPr 1028
      |||
3211 GATGGAAGAAAGATATATGCCAAATA..... 3237
1028 oLeuLeuArgAsnSerLeuAspAspLeuValGlyProProSerAsnLeu 1045
      |||
3238 .....TTAGAAAGCTCAGCT.....ATGAATAGACTCCAGGGAATCAGA 3277
1045 LuGlyGlnSerAsp 1049
3278 AAGCAGCATATGAT 3291

```

seq\_name: /cgn2\_6/ptodata/1/lna/6B\_COMB.seq:US-09-107-847-1

seq\_documentation\_block:

; Sequence 1, Application US/09107847

; Patent No. 6100062

; GENERAL INFORMATION:

; APPLICANT: DUCKWORTH, DAVID

; APPLICANT: MICHALOVICH, DAVID

; TITLE OF INVENTION: NOVEL USE

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSER: Ratner & Prestia

; STREET: P.O. Box 980

; CITY: Valley Forge

; STATE: PA

; COUNTRY: USA

; ZIP: 19482

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/107,847

; FILING DATE: 30-JUN-1998

; CLASSIFICATION:

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: EP 97304996.8

; FILING DATE: 08-JUL-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Prestia, Paul F

; REGISTRATION NUMBER: 23,031

; REFERENCE/DOCKET NUMBER: GH-30003

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 610-407-0700

; TELEFAX: 610-407-0701

; TELEX:

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 5715 base pairs

; TYPE: nucleic acid

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-107-847-1

alignment_scores:
  Quality: 313.00      Length: 1172
  Ratio: 0.607        Gaps: 54
  Percent Similarity: 44.027  Percent Identity: 19.795

alignment_block:
US-09-041-994-2 x US-09-107-847-1 ..

Align seg 1/1 to: US-09-107-847-1 from: 1 to: 5715

34  GluIysArgArgArgGluGlnGluSerLysTyrIleGluGluLeuAlaG1 50
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
378  GAAAGAAACGTAGAGATCAATTTAATGTCATTAAGAAGCTGGATC 427
    50  uLeuIleSerAlaAsnLeuSerAspIleAspAsnIleAsnValLysProA 67
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
428  CATGCTCTCTGTGTAATGCTAGA.....AAGATGG 456

67  sPlyscYsAlaIleLeuLysGluThrValArgIleIleArgGlnIleLys 83
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
457  ACAAACTTACTGTCTGCAGAAAGCATGATTTTACGAAACATGATAA 506
    84  GluGlnGlyLysThrIleSerAsnAspAspValGlnLysAlaAspVal 100
    |||  |||  :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
507  GAA.....ATCAGTCGACAGTCAGATGCTAGTGAAATTCGACAGACTG 550
    100  lSerSerThrGlnGlnGlyValIleAspLysAspSerLeuGlyProLeuL 117
    :|:|  |||  :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
551  GAAACCTTACA.....TTCCTAGTAATGAGAGAGTTTACACAAATTAA 591
    117  euLeuGlnAlaLeuAspGlyPheLeuPheValAlaAsnArgGluAlaAsn 133
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
592  TGTTAGAGCTCTTGATGATGTTTTTTTAGCATCATGACAGATGGAAGC 641
    134  lLeuAlaPheValSerGluAsnValThrGlnTyrLeuGlnTyrLysGlnG1 150
    |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
642  AATAATATGTGTGTGAGAGTGTACTTCACTTGAACAATTTCCATC 691
    150  uAspLeuValAsnThrSerValTyrAsnIleLeuHiscLysGluAspArgL 167
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
692  TGATCTTGTGATCAAACTAATTTAATTTATCCAGAGAGCGAACAAT 741
    167  yAspPheLeuLysAsnLeuProLysSerThrValAsnGlyValSerTrp 183
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
742  CAGAGCTTTATTAATCTCTCTACTCATCTGCTGGAAGTGAATTCATTA 791
    184  ThrAsnGluProGlnArgGlnLysSerHis...ThrPheAsnCysArgMe 199
    |||  |||  :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
792  ACCCGAGAAATTTTAATCAAAAAAATCAGTTAGATTCGTGTGCACAT 841
    199  tLeuMetLysThr.....ProHicAsp..... 206
    |||  |||  :|:|  :|:|  :|:|  :|:|  :|:|  :|:|  :|:|
842  GCTGCGAGACAATAGACCCAAAGAGCCATCTACTATGAATATGTAA 891
    207  .....IleLeuGluAspIleAsnAlaSerProLys 216
    |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
892  AATTATATGAAATTTCAATCTTTAAACAGTATCTCTTCACGACAC 941
    217  .....MetArgGlnArgTyrGluThr 223
    |||  |||  |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
942  AATGGTTTGAAGAACTATACAACGACACATAGCCATCTTATGAAGA 991
    223  tMetGlnCysPheAla.....LeuSerGlnProArgAlaMet 236
    |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
992  TAGAGTTTGTGTTGTAGCTACTGTGAGGTAGCTACACCTCAGTTCATCA 1041
    236  etGluGlnGlyGluAspLeuGlnSerCysMetIleCysValAlaIleArg 252
```

```
1042  AGGAA.....ATCTGC..... 1052
    |||  :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
253  lIleThrThrGlnGluArgThrPheProSerAsnProGluSerPheIleThr 269
    |||  :|:|  :|:|  :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
1053  .....ACTGTTCAAGAACCCCAATGAAGAGTTTACATC 1084
    269  rArgHisAspLeuSerGlyLysValAlaAsnIleAspThrAsnSer... 284
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
1085  TAGACATAGTTTAGAATGGAAGTTTCTGTTCTAGATCAGAGGCGACAC 1134
    285  .....LeuArgSerSerMetArgPro 291
    |||  :|:|  :|:|  :|:|  :|:|  :|:|  :|:|  :|:|
1135  CCATATATAGGTAATTTGCCATTTGAAAGTTCTGGCAACATCA..... 1175
    292  GlyPheGluAspIleIleArgArgCysIleGlnArgPhePheSerLeuAs 308
    |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
1176  GCGTATGAT.....TACTATCATGTGCA 1198
    308  nAspGlyLysSerTrpSerGlnLysArgHisTyrGlnGluAlaTyrLeuA 325
    |||  :|:|  :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
1199  TGACCTAGAAAATTTGCCAAAATGTCATGACACCTTAATGCAATAT... 1244
    325  snGlyHisAlaGlnThrProValTyrArgPheSerLeuAlaAspGlyThr 341
    |||  :|:|  :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
1245  ..GGGAAAGCCAAATCATGTTATTTAGGTTCTGACTAAGGGGCAACAG 1292
    342  lLeuAlaThrAlaGlnThrLysSerLysLeuPheArgAsnProValThrAs 358
    :|:|  :|:|  :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
1293  TGGATTTGGCTTACAGCT..... 1310
    358  nAspArgHisGlyPheValSerThrHisPheLeuGlnArgGluGlnAsnG 375
    |||  :|:|  :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
1311  .....CATTTATATATCACTTACCAT.....CAGTGAAT 1341
    375  lYTyArgProAsnProAsnProValGlyGlnGlyIleArgProMet 391
    :|:  :|:  :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
1342  CA.....AGGCGCAAGCTT 1355
    392  AlaGlyCysAsnSerSerValGlyGlyMetSerMetSerProAsnGlnG1 408
    |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
1356  ATTGTTTGTACACACGTGTAGTAAATGCA.....GA 1390
    408  yLeuGlnMetProSerSerArgAlaTyrGlyLeuAlaAsp.....ProS 423
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
1391  AGTTAGGCGCTGAAGACGACGAGAACTTGGCATTGAAGAGCTCTTCTCG 1440
    423  eThrThrGlyGlnMetSerGlyAlaArgTyrGlyLysSer..... 437
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
1441  AGACAGCTGTGACAAAGC.....CAAGATTCTGGGTGAGATTAATGCT 1484
    438  ..AsnIleAlaSerLeu.....ThrPr 444
    |||  :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
1485  AATAACACAGTCAGCTCAAGAAAGCATTTGAAAGTTTGTATCACAGCC 1534
    444  oGlyProGlyMetGlnSerProSerSerTyrGlnAsnAsnAsnTyrGlyL 461
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
1535  AACCCTTGTGCTCTCTCTCGAGTTCAAGAAATCATCTCACAGCGCG 1584
    461  euAsnMetSerSerProProHicGlySerProGlyLeuAlaProAsn... 476
    :|:  :|:|  :|:|  :|:|  :|:|  :|:|  :|:|  :|:|
1585  TC.....TCAGACCTTCTCTCAACACCAACCAAGATCCCGAGGAT 1635
    477  .....G1 477
    :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:
1626  ACGAGCACTCCACCCAGCAGCATTTACCAGCTCATAGAAAGATGTGCA 1675
    477  nGlnAsnIleMetLysSerProArgAsnArgGlySerProLysIleAlas 494
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
1676  AAGAGGTCATCATTTAGTACTAGTCCATTAATCCAGAGTCTGTGGTT 1725
    494  eRhHisGlnPheSerProValAlaGlyValHisSerProMetAlaSerSer 510
    ||  :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
```



1726	CATCATTAACACAGCAGCAGTATGTCTCAAGCTACAAATTTTACCAATTTTCA	1775
511	gLyasnThgLyAsnHisSerPheSerSerSerSerSerLeuSerAlaLeuG1	527
1776	CAAGGATGTCCCAAGTTTCAGTTT...TCAGTCAATTAAGAGCGCATGCA	1822
527	palalieserGlu.....	531
1823	ACATCTGAAGAAGCAAAATTGGAAACAAGCAGCATGATGATGAACCAATA	1872
531	.....	531
1873	TTTCATCGGCACACAGACAGTAAGAAAAATTCAAGAAACAATTTCAGATG	1922
532	.....gLyValG1YThrSerLeuSerThrLeuSerProGly..	545
1923	GTCCATAGTGCAGGGCGGTGAGATGTTTTTGCACAAATCAAAATCTCGGTT	1972
546	.....ProLyLeuAspAsnSerProAsnMetAsnIleThrG	558
1973	GAATTTTGGTTCGGTTCAACTTTCTTCTGGAAATTCATCTGAACATCCAGC	2022
558	LeuProSerLyAlSerAsnGlnAspSerLySerProLeuGlyPheYr	574
2023	AACTTGCACCTTAAATATGCAAGCCCAAGTTTGTTCCT.....	2060
575	CysAspGlnAspProValGluSerSerMetCysGlnSerAsn...Se	589
2061	.....ACTAACCAAGTTCAAAAGTGGATGATGATCTGGACACATTTGGCAC	2104
589	ratAspHisLeuSerAspLySglSerLySglSerSerValGluGlyA	606
2105	AACTCGACACATGATACAAACACACAGACTTTACAGAGTACA....TCAA	2148
606	lagLysnGlnArgGlyProLeuGluSerLySglYnHisLySglLeuLeu	622
2149	CTCAGAGTAAACAATAATGTACTGAGT.....	2174
623	GlnLeuLeuThrCysSerSerAspArgGlnHisSerSerLeuThrAs	639
2175	.....GGGCACTACGCAAAACATC	2194
639	nsrProLeuAspSerSerCysLySglSerSerValSerValThrSerP	656
2195	TCATACC.....AGTCAGACACAGACACTCTTACAGGCC	2229
656	ro.....SerGlyAlSerSerSerThrSerGlyLyAlSerSer	669
2230	CACGTGTATAACTGTATGTGATTTCTCAGCTCAGCC.....	2267
670	ThrSerAsnMetHisGlySerLeuGlnGlyuShiArgGlyLeuHis	686
2268	.....GGAGCATGTGCTCAG.....ATTTCATC	2290
686	sLySglLeuGlnAsnGlyAsnSerProAlaGlyAlaLeuAlaSerIleThr	702
2291	TAGTATGCCACAAACACAGCACCACAGAGTGTGCAGTACTACATTCACCTC	2340
703	.....AlaGlnAlaThrGlyLysAsp	709
2341	AGCAGCAGCAGATTAAGATTTTCTCAAGGTCAAAACACTGTGACCAAAATTA	2390
710	ThrSerSerIleThrSerCysGlyAsp.....	718
2391	GTGACAGCTCTCTGTAGCTTGTGGCAGTCAATGTACTACTATGCT	2440
719	.....GlyAsnValVal.....LysGlnGng	726
2441	TATGGCCACAGTGTGATGATCATCTACTTTTCTACACACAGCAC	2490
726	LeuSerProLySglYnHisGlnAsnAlaLeuLeuArgTylLeuLeu	742
2491	AG.....	2492

743	AspArgAspAspProSerAspAlaLeuSerTyrGluLeuGlnProGlnVal	759
2493	.....TCAAGCAATCTTCAGTAAACGACGACGACGACCA	2527
759	IgIuGlyValAspAsnLysMetSerGlnCysThrSerSerThrIlePro	776
2528	GCAG.....AGTCCAGGACGACGACGCTCACTTCAGTTACGACCAAT	2571
776	erSerSerGlnIuLysAspProLysIleLysThrGluThrSerGlnIu	792
2572	CTCAGGCTCAGCTGACCCAGCCA.....	2594
793	GlySerGlyAspLeuAspAsnLeuAspAlaIleLeuGlyAspLeuThrSe	809
2594	.....	2594
809	rSerAspPheTyrAsnAsnSerIleSerSerAsnGlySerHisLeuGlyT	826
2594	.....	2594
826	hIryGlnGlnValPheGlnGlyThrAsnSerLeu..GlyLeuLysSer	841
2595	..CCGAAACAATTTTACAGACTTCAGATTGCTCCATGAGGAATCCCA	2642
842	SerGlnSerValGlnSerIleArgProPro..TyrAsnArgAlaValSerI	858
2643	ACTCAACATCATTTCTCTCTGCTGCTGCTCTCTCAACACGACGACCTTCC	2692
858	eu.....AspSerProValSerValGlySerSerProPro	869
2693	TACAGTACATCAACGACCAACATCAGCTTCACGACACGACGACCAACCC	2742
870	ValLysAsnIleSerIlePheProkMetLeuProLysGlnPrometLeuG	886
2743	GGCACAGGACTGACACTTGCCCGACCTCCAAAG.....	2778
886	gLyIAsnProArgMetLeuMetAspSerGlnLysAsnTyrGlySerSerMetG	903
2779	.....TTCAACACAGTACGACACGCTGCTTCCTCTCTT	2811
903	LysGlyProAsnArgAsnValThrValThrGlnThrProSerSerGlyAsp	919
2811	.....	2811
920	TryGlyLeuProAsnSerLysAlaGlyArgMetGluPrometAsnSerAs	936
2812	.....GACATCAAGGAGGAAAGGATGGCCCATTTAAGAT..	2847
936	nSerMetGlyArgProGlyGlyAspTyrAsnThrSerLeuProArgProA	953
2848	.....TACTCAGATGACCTGAGGAAA....	2868
953	lAlaLeuGlySerIleProThrLeu.....ProLeuArgSerAsn	966
2869	.....GGAGGAAAGTTCAGCAGATTTCATGACGATCGATGTGAGGTT	2913
967	SerIleProGly.....AlaArgProValLeuGlnGlnGlnGlnG	980
2914	CTAGTTCTCGAATTAGTTGGGACAGAAATGCTCCTAAGTCATACAGAT	2963
980	nMetLeuGlnMetArgProGlyGluIleProMetGlyMetGlyAlaAsn	997
2964	GTACATTAATACACGCCAGCA.....GGAGGTATC	2995
997	roTyrGlyGlnAlaAlaAlaSerAsnGlnLeuGlySerTyrPro.....	1011
2996	ATAGGGGACACAGCCAGTTTGACAGTGTTTAAAGTGCTGATATTTT	3045
1012	AspGlyMetLeuSerMetGlnValSerHisGlyThrGlnAsnArgPr	1028
3046	GATGGAAAGAAATATATGCAAAAT.....	3072

```

seq_name: /cgn2_6/pdata/L/ina/5B.COMB.seq:US-08-816-693A-1
seq_documentation_block:
    : Sequence 1, Application US/08816693A
    : Patent No. 5874241
    : GENERAL INFORMATION:
        : APPLICANT: Takahashi, Joseph S
        : APPLICANT: Turek, Fred W
        : APPLICANT: Pinto, Lawrence H
        : TITLE OF INVENTION: Clock Gene and Gene Product
        : NUMBER OF SEQUENCES: 53
        : CORRESPONDENCE ADDRESS:
        : ADDRESSEE: Dressler, Rocky, Milnamow & Katz
        : STREET: Two Prudential Plaza, Suite 4700
        : CITY: Chicago
        : STATE: Illinois
        : COUNTRY: USA
        : ZIP: 60601
        : COMPUTER READABLE FORM:
        : MEDIUM TYPE: Floppy disk
        : OPERATING SYSTEM: IBM PC compatible
        : SOFTWARE: Patent Release #1.0, Version #1.30
        : CURRENT APPLICATION DATA:
        : APPLICATION NUMBER: US/08/816,693A
        : FILING DATE:
        : CLASSIFICATION: 435
        : ATTORNEY/AGENT INFORMATION:
        : NAME: No. 5874241Ihrup, Thomas E
        : REGISTRATION NUMBER: 33,268
        : TELECOMMUNICATION INFORMATION:
        : TELEPHONE: 312-616-5400
        : TELEFAX: 312-616-5460
        : INFORMATION FOR SEQ ID NO: 1:
        : SEQUENCE CHARACTERISTICS:
        : LENGTH: 7498 base pairs
        : TYPE: nucleic acid
        : STRANDEDNESS: single
        : TOPOLOGY: linear
        : MOLECULE TYPE: DNA (genomic)
        : FEATURE:
        : NAME/KEY: CDS
        : LOCATION: 389..2954
        : US-08-816-693A-1
alignment_scores:
    Quality: 309.50      Length: 1138
    Ratio: 0.614         Gaps: 51
    Percent Similarity: 44.288   Percent Identity: 19.947
alignment_block:
    US-09-041-994-2 x US-08-816-693A-1 ..
Align seg 1/1 to: US-08-816-693A-1 from: 1 to: 7498
34 GtUlyArGArgArGgUGngInGUSeLySTyTlGnGUleuAlaGl 50
|||||:::|||||:::|||::| | |||||::|::|
515 GAAMAAGAAACGTAGATCAGTCATGCTCATTAAGAGCGTGCGGC 564
|||||:::|||||:::|||::| | |||||::|::|
50 UeuIUeSerAlaAsnuLeuSerSplEaspaNpheASvAlLysPro 67
|||||:::|||||:::|||::| | |||||::|::|
565 TAGCTTCCGTGGTAGCCGAGA.....AAGATGG 593
67 splUcySaLaIaIeUlySglUThrValaRgInIleaRgInIleYs 83

```

[illegible]

1382 ..GGAAAGCAATCGTGTACTATGATTCTGACCAAGGCCACAG 1429  
 342 ILevalThrIaGlnThrLysSerLysLeuPheArgAsnProValThrAs 358  
 1430 TGGATATGGCTTCAGACT..... 1447  
 358 nAspArgnHISglYpHevalSerThrHISpHeuGlnArgGluGlnAsnG 375  
 1448 .....CATTATTATTTACTTACCAT.....CAGGGCAATT 1478  
 375 LylYrArgProAsnProAsnProValGlyGlnGlyLeuArgProPmet 391  
 1479 CA.....AGCCGAGACTTC 1492  
 392 AlAGlyLysAsnSerSerValGlyGlyMetSerMetSerProAsnGlnG 408  
 1493 ATGTGTTGTACTCACACTGTAGTAAGTTATGA.....GA 1527  
 408 YLeuGlnMetProSerSerArgAlaTyrgLYLeuAlaAsp.....ProS 423  
 1528 AGTTAGGGCTGAAAGACGGCGAGACTTGGCATTGGAAGACTCTCTCTG 1577  
 423 eThrThrIaGlnMetSerGlyAlaArgTyrgLYSerSer..... 437  
 1578 AGACAGCTGCTGACAAAGC.....CAAGATTCTGGTCTGACAAATCGT 1621  
 438 ...AsnIleAlaSerLeu.....ThrPr 444  
 1622 ATCAGACAGACTGAGTCTCAAGAGACACTGMAAGTTTATACACACCC 1671  
 444 OGlyProGlyMetGlnSerProSerSerTyrgLInAsnAsnAsnTyrgLY 461  
 1672 AACTCTTCTGCTCTCTAGAAAGCTCAGAAAGCTCATCTCACACCCGAG 1721  
 461 euAsnMetSerSerProHISglYSerProGlyLeuAlaProAsnGln 477  
 1722 TC.....TCAGACCTTCTCTCCACACGCAAGATCCCTACTGAT 1762  
 478 GlnAsnIleMetIleSerProArgAsnArgGlySerProLysIleAlaSe 494  
 1763 ACTAGC.....ACTCTCCACAGCAG..... 1783  
 494 rHISGlnPheSerProValAlaGlyValHISserProMetAlaSerSerG 511  
 1784 .....CATTTGCCAGCTCATGAAAGA 1805  
 511 LysAnThrIaGlnHISserPheSerSerSerSerLeuSerAlaLeuGln 527  
 1806 TGACACAGCGGAGGTGCTCTTCAGCAGTCAGTCATTAAC..... 1846  
 528 AlAlaIleSerGluGlyAlaGlyThrSerLeuLeuSerThrLeuSerSerPr 544  
 1847 .....TCCAGTCAGTTGGTTCATCATTAACAG..... 1876  
 544 OGlyProLysLeuAspAsnSerProAsnMetAlaIleThrGlnProSerL 561  
 1877 .....CCAGCATGTCTCAAGCTGCAAAATTACCAATT.....CCACAAG 1916  
 561 yValSerAsnGlnAspSerSerLysSerProLeuGlyPheTyrgYAspGln 577  
 1917 GCATGTACAGATTTCAGTTTTCAGCTCAGTTAGA..... 1951  
 578 AsnProValIleuSerSerMetCysGlnSerAsnSerArgAspHISLeuSe 594  
 1952 .....GCCATGACAGCATTTTAA 1968  
 594 rAspLysGlnSerLysGlnSerSerValGluGlyAlaGlnAsnGlnArg 611  
 1969 AGACACAGCTAGAGCAGCGACACGATGATAGAGCAAAATTTTCATGGGC 2018  
 611 LylProLeuGlnSerLysGlyHISLysLysLeuGlnLeuLeuThrCys 627  
 2019 AGCAAGAAAGACTAAAGAAATTTCAAGACAACTTCAGATG..... 2059

628 SerSerAspAspArgGlyHISSerSerLeuThrAsnSerProLeuAspSe 644  
 2059 ..... 2059  
 644 rSerCysLysGlnSerSerValSerValThrSerProSerGlyValSerS 661  
 2059 ..... 2059  
 661 eSerThrSerGlyGlyValSerSerThrSerAsnMetHISGlySerLeu 677  
 2060 .....GTCATGTGCAGCAGG 2074  
 678 LeuGlnGlnLysHISArgIleLeuHISLysLeuLeuGlnAsnGlyAsnSe 694  
 2075 CTACAG.....ATGTTTTGCAGCAATCCAAC.. 2101  
 694 rProAlaGluValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrS 711  
 2102 .....CCGTGATGTGAATTTG 2117  
 711 eSerIle...ThrSerCysGlyAspGlyAsnValValLysGlnGlnGln 726  
 2118 GTTCTGTCAACTTCTCTCGAAATTCATAATC.....CAGCAG 2158  
 727 LeuSerProLysLysLysGlnAsnAsnAlaLeuLeuArgTyrgLeuLeuAs 743  
 2159 CTCACACT..... 2167  
 743 pArgAspAspProSerAspAlaLeuSerLysGluLeuGlnProGlnValG 760  
 2168 .....GTAAATATCAAGCCAGGTT. 2188  
 760 LysGlyValAspAsnLysMetSerGlnCysThrSerSerThrIleProSer 776  
 2189 .....GTGCCCTGCT 2197  
 777 SerSerGlnGlnLysAspProLysIleLysThrGluThrSerGluGlnG 793  
 2198 AACCGAGTTTCAG..... 2209  
 793 ySerGlyAspLeuAspAsnLeuAspAlaIleLeuGlyAspLeuThrSers 810  
 2210 ..AGTGCATATTC..... 2221  
 810 eAspPheTyrgAsnAsnSerIleSerSerAsnGlySerHISLeuGlyThr 826  
 2222 .....AGCACAGCCACGACATG..ATA 2242  
 827 LysGlnGlnValPheGlnGlyThrAsnSerLeuGlyLeuLysSerSerG 843  
 2243 CAGCAACAGACTTTTACAAGTATCATCA.....ACTCGCA 2277  
 843 nSerValGlnSerIleArgProProTyrgAsnArgAlaValSerLeuAspS 860  
 2278 GAGTCAACAAGTGTATAGATGAGTCAGACACAGTCAACAGCTCTTCCAA 2327  
 2328 GTCAGACACGAGCAGCTCAGACAGCCCACTGTACAT..... 2365  
 877 ProMetLeuProLysGlnProMetLeuGlyGlyAsnProArgMetMetAs 893  
 2366 ACGATGTGATTTCCTCAGCTGCAGCTGGG..... 2395  
 893 pSerGlnGlnAsnTyrgLYSerSerMetGlyGlyProAsnArgAsnValT 910  
 2396 .AGCATGTCTCAAGTTTCATCCAGTATG...CCACAGAAAGTACCCAGA 2441  
 910 hValThrGlnThrProSerSerGlyAspTyrgLYLeuProAsnSerLys 926  
 2442 GTGTACAGTACACTACGTCTACTCAGACAGACAGATTAAGATTTTTCAA 2491

```

927  AAGlAGrAgtmetluprPmetlAnSerAAnSermetGlYnPrProGlYgl 943
    :::      ::      ::      ::      ::      ::      |||
2492  GGT.....CAGCACTTGTGGACCAATTAGTACTGCTCTGTAGC 2532
    :
943  yAsPrYrAsnThSerleuPrArGrPrAlaLeuGlYSerle... 958
    :      ::|||      |||      |||
2533  TTGTGGGGCGCGTATGGTACAGTACACATGACATGATGGTCAGGTGGTGA 2582
    :
959  .....ProThleuPrOleuAArgSerAnSerleIleProGlYAlaArg 972
    |||||      ::      ::      ::      ::      ::
2583  CTGCGCATCTTACCTTCCGCCACACACAGGACAGGCGCACAG...ACATTGA 2629
    :
973  ProValleuGlInglInglInglInglInglInglInglInglInglIngl 989
    |||      |||||      |||||      |||      ::      ::
2630  TCGGTAAACACACAGACAGACAGACAGACAGACAGACAGACAGACAGAACAA 2679
    :
989  ePrOmetGlYmeGlYAlaAn.....ProTylGlylnA 1001
    ::      ::      ::      ::      ::      |||      |||
2680  GCAACAACAACACACAGAGTGTCCAGGAACAGACAGACAGCTTCTTCAGTTGAC 2729
    :
1001  lAlAlaIaSerAnGlInglInglYSerTrPrOArSrlGlymetLeuSermet 1017
    |||||      |||||      |||||      |||||      ::      ::
2730  AGCCAGCTGACGGCCACAGCTGGCCAGCCACACAGACAGCTTCTTACACACA 2779
    :
1018  GlUglInValSerHlsglYnThrGlInAsnArgPrleuLeuArgSer.. 1033
    ::      ::      |||||      ::      |||      ::      ::
2780  TCTAGTGTGCTCCACGGGAATCCTTCGCACACAGCTCAATCCTCTGCTGC 2829
    :
1034  .....LeuAsPrleuValGlYPrOProSerAn..... 1043
    |||||      ::      ::      ::      ::      ::
2830  CTTCGACATGACACAGACAGCTTCTCTCTGTCACACACAGACACACC 2879
    :
1044  .....LeuGlInglInSer 1048
    :
2880  AGCCTCAGCAGCAACAGACAGCTTCTGCGACAGAGACTGCAGCTGACT 2929
    :
1049  AArgGlAArgAlaLeuLeuArgInleuHlsth...LeuLeuSerAn.T 1064
    |||      ::      |||      |||||      ::|||
2930  GACCTTCCAGGTCCAGGCCACAGTACGACACACACTTCTCTGTGCAT 2979
    :
1064  hAsPrAlaThrGlYleuGlInglInleAsPrArgAlaLeuGlYlePrGlIn 1080
    ::      |||||      ::      ::      |||||      ::
2980  GCGAGAGGAAGGGGATGGCCGACGAAGAATGCTGATGGCATCGGTGCA 3029
    :
1081  LeuValAsnGlIn 1084
    |||      ::|||
3030  GAAGTGAACAG 3041

seq_name: /cgn2_6/ptodata/1/lna/6A_COMB.seq:US-08-885-291-1

seq_documentation_block:
Sequence 1, Application US/08885291A
Patent No. 6057125
GENERAL INFORMATION:
APPLICANT: Takahashi, Joseph S.
APPLICANT: Turek, Fred W.
APPLICANT: Pinto, Lawrence H.
TITLE OF INVENTION: CLOCK GENE AND GENE PRODUCT
FILE REFERENCE: 0290-5
CURRENT APPLICATION NUMBER: US/08/885,291A
CURRENT FILING DATE: 1997-06-30
EARLIER APPLICATION NUMBER: 08/816,693
EARLIER FILING DATE: 1997-03-13
NUMBER OF SEQ ID NOS: 55
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 7498
TYPE: DNA
ORGANISM: Mus musculus
US-08-885-291-1
alignment_scores:

```

Quality:	309.50	Length:	1138
Ratio:	0.614	Gaps:	51
Percent Similarity:	44.288	Percent Identity:	19.947
<p>alignment_block:</p> <p>US-09-041-994-2 x US-08-885-291-1 ..</p>			
<p>Align seg 1/1 to: US-08-885-291-1 from: 1 to: 7498</p>			
34	GLULysArGArGArGluGluGluSerLysTrIleGluGluLeuAlaGlu	50	
515	GAAlAAAGAAAGCGTAGAGATCAATTCATAGCTCTCTTAAGAGCGGGGCTC	564	
50	uLeuIleSerAlaIAsnLeuSerSPlleAsnPrheAsnValLysPro	67	
565	TAAGCTCTCTGGTAAGCCGAGA.....AAGATGG	593	
67	sPLysAlaIleLeuLysGluThrValArgInIleArgInIle	83	
594	ACAAAGCTACTGCTGTACAGAAAGCATGTGATTTTGGCCAAACATATA	643	
84	GLuGluGluLysThrIleSerAsnAspAspValGluInLysAlaAspVal	100	
644	GAA.....ACCCATGCAAGATGAGATGATTCACAGAGCATG	687	
100	IleSerThrGluGluValIleAspLysAspSerLeuGlyProLeu	117	
688	GAAGCCACAC.....TTCCATGATTAAGAGCTTTCACAGTTAA	728	
117	euleuGlnAlaLeuAspGlyPheLeuPheValValAsnArgGluAlaAsn	133	
729	TGTTAAGAGCTCTGTAGTGTATTTTATAGCATCATGACAGATGAAGAT	778	
134	IleValPheValSerGluAsnValThrGluTyrLeuGluInLysGluGlu	150	
779	ATAATATATGATCTGAGAGCTGTAACCTGCTTCAATTAACCTTACATC	828	
150	uAspLeuValAsnThrSerValTyrAsnIleLeuInGluGluAspArgL	167	
829	TGATCTGTGGATCAAGATATATTATTATTTATCCSAGAGSAGAACTT	878	
167	ysAspPheLeuLysAsnLeuProLysSerThrValAsnGluValSerTrp	183	
879	CAGAGGTTTATAAGATACTCTACTCATCTGCGGAAGAGCATCATTA	928	
184	ThrAsnGluProGlnArgInLysSerHis...ThrPheAsnCysArgMet	199	
929	ACCCCTGAGTACTAAATCAAAAAACAGTTAGATAATCTGTGTGCAT	978	
199	tleuMetLysThr.....ProHisAsp.....	206	
979	GCTTCGAGAACAAATGAGCCAAAGAGCCATCCACCTATGATATATGGA	1028	
207	.....IleLeuGluAspIleAsnAlaSerProGlu	216	
1029	GATTTATAGCAAAATTTTAAATCTTTTAAACAGTGTATCAACTTCAACACAC	1078	
217	.....MetArgGlnArgTyrGluThr	223	
1079	AATGTTTGAAGAACTATACAAAGCACACATAGGCTCTTATGGAAGA	1128	
223	rMetGlnCysPheAla.....LeuSerGlnProArgAlaMet	236	
1129	TAGAGTTGTTGTGTAGCTACTGTCAGATTAGTACACACCTAGTTACATCA	1178	
236	eGluGluGluGlyGluAspLeuGlnSerCysMetIleCysValAlaArgArg	252	
1179	AGGAA.....ATGTGT.....	1189	
253	IleThrThrGlyGluArgThrPheProSerAsnProLysSerPheIleThr	269	
1190	.....ACTGTGAAGAACCAATATGAAAGGTTTACATC	1221	

```

1877 ... CCAGGAGTCTCAAGCTGCAAAATTATTACCAATT ..... CCACAG 1916
561 ysa1serasnclnaspserltysserproleuglyphetlycysapln 577
1917 GCATGTCACAGTTTCAGTTTCAAGTCAAGTTAGGA ..... 1951
578 AsnProValGluSerSerMetCysGlnSerAsnSerArgAspHisLeu 594
1952 ..... GCCATGCACAGACTCTAA 1968
594 rasplysgluserlysgluserServalGluValaGluasnGlnarg 611
1969 AGACCCAGCTAGAGCAGCCGACACGATGATGACGCAAAATTATTCGCGC 2018
611 lProleuGluSerlysglyHisLysLysLeuGlnLeuLeuthrCys 627
2019 AGCAGAAACCTAAGAAATTCAGAGCAACTTCAGATG ..... 2059
628 SerSerAspArgGlyHisSerSerLeuthrAsnSerProLeuAsp 644
2059 ..... 2059
644 rSerCysLysGluSerServalServalThrSerProSerGlyValSers 661
2059 ..... 2059
661 eSerThrSerGlyGlyValSerSerThrSerAsnMetHisGlySerleu 677
2060 ..... GTCCATGCTGCACAGG 2074
678 LeuGlnGluLysHisArgIleLeuHisLysLeuGlnAsnGlyAsnSe 694
2075 CTRACAG ..... ATGTTTGGCAGCAATCAAC. 2100
694 rProAlaGluValAlaLysIleThrAlaGlnAlaThrGlyLysAspHis 711
2102 ..... CCTGATTAATTTGG 2117
711 eSerIle...ThrSerCysGlyAspGlyAsnValValLysGlnGln 726
2118 GTTCCTGTTAACTTCCTCTGGAAATTCCTAAATATC ..... CACAG 2158
727 LeuSerProLysLysGluLysAsnAsnAlaLeuLeuArgTyrLeuLeuAs 743
2159 CTCACACT ..... 2167
743 pArgAspAspProSerAspAlaLeuSerLysGluLeuGlnProGlnValG 760
2168 ..... GTAAATATGCAAGGCCAGGTT. 2188
760 LucIlyAlaAspAsnLysMetSerGlnCysThrSerSerThrIleProSer 776
2189 ..... GTCCCTGCT 2197
777 SerSerGlnGluLysAspProLysIleLysThrGluThrSerGluGlu 793
2198 AACCAAGTTCAG ..... 2209
793 ySerGlyAspLeuAspAsnLeuAspAlaIleLeuGlyAspLeuThrSers 810
2210 .AGTGACATATC ..... 2221
810 eAspPheTyrAsnAsnSerIleSerSerAsnGlySerHisLeuGlyThr 826
2222 ..... ACCACAGCCACACATG. ATA 2242
827 LysGlnGluValPheGlnGlyThrAsnSerLeuGlyLeuLysSerSerG 843
2243 CAGCAACACACTTTCACAAAGTACATCA ..... ACTCAGCA 2277
843 nSerValGlnSerIleArgProProTyrAsnArgValaSerLeuAspS 860

```

```

2278 GAGTCAACAGAGTGTATGATGTCGACACAGTCAAGACGCTCTTCCAA 2327
860 erProvalSerValGlySerSerProProValLysAsnIleSerAlaphe 876
|||::: |||::: |||
2328 GTCACAGACCGAGCACTCTCAGACAGCCACCTGTACAT..... 2365
877 ProMetLeuProLysGlnProMetLeuGlyLysProArgMetMetas 893
|||::: |||::: |||
2366 ACGATGGGTATTCACAGCTGACCTGGG..... 2395
893 pSerGlnLysAsnTyrgLysSerSerMetGlyLysProAsnArgAsnValT 910
|||::: |||::: |||
2396 .AGCATGGTCGACGATTCATCCAGTATG...CCACAGACAGTACCCAGA 2441
910 hValThrcInThrProSerSerGlyAspTyrPylLeuProAsnSerLys 926
|||::: |||::: |||
2442 GTCGACAGTACTACTAGTCTACTCAGACAGACAGATAGATTCTCTCA 2491
927 AlaGlyArgMetGluProMetAsnSerAsnSerMetGlyArgProGlyG 943
|||::: |||::: |||
2492 GCT.....CAGCAACTGTGTGACCAAAATTAGTACTGCTCTGTAC 2532
943 yAspTyrAsnThrSerLeuProArgProAlaLeuGlyGlySerIle... 958
|||::: |||::: |||
2533 TTGTGGGGCCGTCATGTTACCAAGTACCTGCTTATGGTCAAGTGTGA 2582
959 .....ProThrLeuProLeuArgSerAsnSerIleProGlyAlaArg 972
|||::: |||::: |||
2583 CTGCGTATCTTACCTTCCGACACCAACAGCAGCAGCAG...ACATTA 2629
973 ProValLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 989
|||::: |||::: |||
2630 TCGGTAAACACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 2679
989 eProMetGlyMetGlyAlaAsn.....ProTyrGlyGlnA 1001
|||::: |||::: |||
2680 GCAACACACACAGCAGAGTTCACAGACAGACAGCAGCTTCTTACGTTCA 2729
1001 lAlaAlaSerAsnGlnLeuGlySerTyrProAspLysMetLeuSerMet 1017
|||::: |||::: |||
2730 ACCCAGCTCAGGCGCGAGTGGCGCAGCAGCAGCAGCAGCAGCAGCAGCA 2779
1018 GlnGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSer.. 1033
|||::: |||::: |||
2780 TCTAGTGTCTCCACGCGGATCTTCGACACAGCTCATCTCTGCTGCTGC 2829
1034 .....LeuAspAspLeuValGlyProProSerAsn..... 1043
|||::: |||::: |||
2830 CTTTCCACTACACAGCAGCAGCTTCTCTCTGCGCACACAGCAGCAGCACC 2879
1044 .....LeuGlnGlyGlnSer 1048
|||::: |||::: |||
2880 ACCCTCAGCAGCAGCAGCAGCTTCTCGGACAGCAGCAGCAGCAGCTGAT 2929
1049 AspGlnArgAlaLeuLeuAspGlnLeuHisThr...LeuLeuSerAsn..T 1064
|||::: |||::: |||
2930 GACCTTCCAGGTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGT 2979
1064 hAspAlaThrGlyLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1080
|||::: |||::: |||
2980 GCGAAGAGCAAGGAGGATGGCCAGAAAGAAATCGCTGATGGCGGTCA 3029
1081 LeuValAsnGln 1084
|||::: |||::: |||
3030 GAAGTTGAACAG 3041

```

seq\_name: /cgr2\_6/ptodate/1/ina/5A\_COMB.seq.us-08-227-536-1

```

seq_documentation_block:
; Sequence 1, Application US/08227536
; Patent No. 5658784
; GENERAL INFORMATION:
; APPLICANT: Eckner, Richard

```

```

; APPLICANT: Ewen, Mark
; APPLICANT: Livingston, David
; TITLE OF INVENTION: NUCLEIC ACID, ENCODING TRANSCRIPTION
; TITLE OF INVENTION: FACTOR P300 AND USES OF P300
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
; STREET: Ten Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/227,536
; FILING DATE: 14-Apr-1994
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Ph.D., Kathleen A.
; REGISTRATION NUMBER: 34,380
; REFERENCE/DOCKET NUMBER: DECT-308XX
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 451-0313
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9046 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1200..8441
; US-08-227-536-1

alignment_scores:
Quality: 274.00 Length: 699
Ratio: 0.828 Gaps: 29
Percent Similarity: 47.353 Percent Identity: 23.176

alignment_block:
US-09-041-994-2 x US-08-227-536-1 ..
Align seg 1/1 to: US-08-227-536-1 from: 1 to: 9046

770 ThrSerSerThrLeuProSerSerSerGlnGlnLysAspProLysIleLys 786
|||::: |||::: |||
6562 TCTGCTGATCATCATGCAAGCAGCTCCAGAGAAACAAATGCGCGTGGCG 6611
786 sThrGln...ThrSerGlnGlnGlySerGlyAspLeuAspAsnLeuAspA 802
|||::: |||::: |||
6612 TTTCCTCCTAACATCAACGAGAAAGCTCCGCG..... 6642
802 lAlaLeuGlyAspLeuThrSerSerAspPheTyrAsnAsnSerIleSer 818
|||::: |||::: |||
6643 .....AGCAACAGCTGACAGCAGCAGC...TACAGCAGGCCCAATGC 6681
819 SerAsnGlySerHisLeuGlyThrLysGlnGlnValPheGlnGlyThrAs 835
|||::: |||::: |||
6682 TTCCGACGAGGATGGCCAGACATGACGCGAGCTGTGTGGTGGCAGCA 6731
835 nSerLeuGlyLeuLysSerSerGlnSerValGln.....S 847
|||::: |||::: |||
6732 CAGGGCTTCCTCCCGCCACCTCTGCGCACATCCACAGACCAACAACTGGCCA 6781

```

```

847 erileargProProTyraSerArg.....AlaValSerLeuAspSer 860
      ::::::::::::::::::::|||
6782 ACAGCCACCAACCCCGACAGCCGACCCACCTTCTCAGCCTTCAGCCTA 6831
      |||
861 ProValSerValGlySerSerProPro.VallyAsnIleSerAlaIlePhe 877
      |||:::|||||
6832 CCCCCTCCCAATGAGCATGCCACCCCTACTTGGCCAGACTCAAGCTGCGGC 6881
      |||:::|||||
877 rometLeuProLysGlnProMetLeuGlyGlnProArgMetMetAsp 893
      |||:::|||||
6882 CCGTGTGTCCCA..... 6892
894 SerGlnGluAsnTyGlySerSerMetGlyGlyProAsnArgAsnValTh 910
      ::::::::::::::::::::|||
6893 .....GGGTAGGCACAGCCGCGGAGTGCCTTC..... 6922
910 rValThGlnThrProSerSerGlyAspTPTGlyLeuProAsnSerIysA 927
      ::::::::::::::::::::|||
6923 .....AACCCCTCTCA...GACTGCTCAGCCACCCCTTCAGGG 6959
927 laGlyArgMetGluProMetAsnSerAsnSerMetGlyArgProGlyGly 943
      |||:::|||||
6960 CCCCACCTACAGCATGTGAATGGCAATGCAGATTCAGAGAGCAGCGGA 7009
      |||:::|||||
944 AspTyraSerThrSerLeuProArgProAlaLeuGlyGlySerIleProTh 960
      |||:::|||||
7010 GAGCGACGCCGACGATGGCCCGACGTCGAATTTTCAAGGCCAATTCAC 7059
      |||
960 r.....LeuProLeuArgSerAsnSerIleProGlyAlaArgProVal 975
      |||:::|||||
7060 ACCGATGAGCCCGGATGATCCATGCGCCCGCATGGGTATGAAACCCACCT 7109
975 eugGlnGlnGlnGlnMetLeu.GlnMetArgProGlyGluIleProme 991
      ::::::::::::::::::::|||
7110 CCATGACCGAGAGTCCCGATTTGGAGGAGCGAGATGGAGCCGAC 7159
991 tGlyMetGlyAlaAsnProTyGlyGlnAlaAlaIleSerAsnGlnLeu 1008
      |||:::|||||
7160 AGGGATTCACCAACAGCCA..... 7178
1008 IySerTTPProAspGlyMetLeuSerMetGlnGlnValSerHisGlyThr 1024
      |||:::|||||
7179 ..CCCTGAGCCCAAGAGATGGCTTCCATGAGCCCGCAACTAGACTGGG 7226
1025 GlnAsnArgProLeuLeuArgAsnSerLeuAspAspLeuValGlyProPr 1041
      ::::::::::::::::::::|||
7227 ATGCCAAGGCCAGCATGATGATGAGTGGCCAGCAT.....GGTCAACC 7270
1041 oSerAsnLeu..... 1044
      |||:::|||||
7271 TTTCACATATGGCTCCACACAGGATTTGGCCAGTATGATCCAGCCAC 7320
1045 .....GlnGlyGlnSerAspGlyArgAlaLeuLeuAspGlnLeuHisThr 1059
      |||:::|||||
7321 TCAAAACGAGGACACTGTCTCACAACACCTTCAAAACCTTTTGGGACT 7370
1060 LeuLeuSerAsnThrAspAlaThrGlyLeuGlnGluIleAspArgAlaLe 1076
      |||:::|||||
7371 CTGAGGTCTCCAGCTCTCCCTGCGACACACAGAGGTCTTATATCTT 7420
1076 u.....GlyIleProGluLeuValAsnGlnGlyGlnAlaLeuGluProLysG 1092
      |||:::|||||
7421 TCACGCCACCCCGAGCTGTG..... 7442
1092 InAspAlaPheGlnGlyGlnGlnAlaAlaValMetAspGlnIysAla 1108
      |||:::|||||
7443 .....GCTGATTCATCAAGCAGCGGCT 7466
1109 GlyLeuTyGlnGlnThrTyProAlaGlnGlyPro..... 1120
      ::::::::::::::::::::|||
7467 GCCAAGTATGCAACTCTAATCTCAACAACCATCCCTGCGCAGCTGGCAT 7516

```

```

1121 .....ProMetGlnGlyG 1125
7517 GCCCCAGGGGAGCCAGGCGCTACAGCCACCTACCATCCAGGTGAGAGG 7566
1125 LysPheHisLeuGlnGlyGlnSerProSerPheAsnSerMet..... 1138
      |||:::|||||
7567 GGGTCCACTCCATTCAGGCATGCAAGACATGATCCAAATGCAGAGCGCGG 7616
1139 .....MetAsnGlnMetAsnGlnGlnIysAsnPh 1148
      ::::::::::::::::::::|||
7617 GTTCAGAGGCGTGGCTGCCCGCCAGCACAACACAGCAGCAGCTCCAGCC 7666
1148 eProLeuGlnGlyMetHisProArgAlaAsnIleMetArgProArgThra 1165
      |||:::|||||
7667 ACCCATGGAGGAGTGAAGCCCGAGCTCAGCATATCAATGATGAACACA 7716
1165 snThr.....ProLysGlnLeuArgMetGlnLeuGlnIleAsnGlyGlnGly 1180
      |||:::|||||
7717 ACACCATGCGCTTCACATTCGA.....GACATCTTGAGACGA 7754
1181 GlnGlnPheLeuAsnGlnSerArgGlnAlaLeuGluLeuIleGlyshetGluAs 1197
      |||:::|||||
7755 CAGCAATGATGACAAACAGCAGACAGCA.....CAGGAGCAGG 7792
1197 nProThrAlaGlyGlyAlaAlaValMetArgProMetMetGlnProGln 1214
      |||:::|||||
7793 GCCAGCAATGAGCCCTGGATGGCAACCAATTAACCATTCAGCAACCCC 7842
1214 lnglyPheLeuAsnAlaGlnMetValAlaGlnArgSerArgGluLeu 1230
      |||:::|||||
7843 AAGAGTGGCTACCCACACAGCCGACGAGCG.....ATG 7880
1231 SerHisHisPheArgGlnGlnArgValAlaMetMetMetGlnGlnGln 1247
      |||:::|||||
7881 CAGATCATCATGCAACAGATGCAAAAGAAATATGGACAGATAGGCA 7930
1247 ngGlnGln.....GlnGlnGln 1254
      |||
7931 GCTTCCCGCAGGCTTGGAGACAGGAGGTGCCAGTCTCAGGCGCTATC 7980
1254 lngGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnAla 1270
      |||:::|||||
7981 AGCAGCAGCTCTTCACACACAGATGGGCTCCCTGTTCAGCCCAACCC 8030
1271 PheSerProPro.....ProAsnValThrAlaSerProSerPhe 1283
      |||:::|||||
8031 ATGAGCCCGCAGCAGCATATGCTCCAAATCAGGCCACGTCACACCT 8080
1283 tAspGlyLeuLeuAlaGlyProThrMetProGlnAlaProProGlnGlnP 1300
      |||:::|||||
8081 ACAAGGC.....CAGCAGA 8094
1300 heProTyGln...ProAsnTyGlyMetGlyGlnGlnProAspProAla 1315
      |||:::|||||
8095 TCCCTATATCTCTCTCCATCAATCAAGTGGCGCTGCCAGCTGTGCTCTT 8144
1316 PheGlyArgValSerSerProProAsnAlaMetMetSerArgMetGln 1332
      |||:::|||||
8145 CCAGCGCCACAGTCCAGCCCGCCACTCCAGTCCCTCCCAAGGATGA 8194
1332 yProSerGlnAsnProMetMetGlnHisProGlnAlaAlaSerIleTy 1348
      |||:::|||||
8195 GCTTCAGCTTCTTCACACCAACGATTTCCACAGCAAGTTCCCCACATC 8244
1349 .....GlnSerSerGlnMetLysGlyTyProSerGly 1359
      |||:::|||||
8245 CTGAGCTGTAGTGCAGCCAGGCAACCCCATGGAACAAGGCAFTTTGCC 8294
1360 AsnLeuAlaArgAsnSerSerPheSerGlnGlnGlnPheAlaHisGlnGln 1376
      ::::::::::::::::::::|||
8295 AGCCCGGACCAATTCATCAATGCTTTCACGCTT.....GCTAG 8332
1376 yAsnProAlaValTySerMetValHisMetAsnGlySerSer 1390

```

```
seq_name: /cgr2_6/prodata/1/lna/PCTUS_COMB.seq:PCT-US95-04682-1
8333 CAATCCA.....GGCATGGCAAACTCCATGTCGTAACG 8366
:||||| :|||||:|||||:|||||:|||||
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: NUCLEIC ACID ENCODING TRANSCRIPTION
TITLE OF INVENTION: FACTOR P300 AND USES OF P300
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
STREET: Ten Post Office Square
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04682
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/227,536
FILING DATE: 14-April-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Holliday C. Heine, Ph.D.
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: DECI-308X999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-2290
FAX: (617) 451-0313
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9046 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1200..8441
PCT-US95-04682-1

alignment_scores:
Quality: 274.00 Length: 699
Ratio: 0.828 Gaps: 29
Percent Similarity: 47.353 Percent Identity: 23.176

alignment_block:
US-09-041-994-2 x PCT-US95-04682-1 ..
Align seg 1/1 to: PCT-US95-04682-1 from: 1 to: 9046

770 ThrSerSerThrIleProSerSerSerGlnGluTyrAspProIleLeu 786
:|||||:|||||:|||||:|||||:|||||
6562 TCTGTCGTCACATGCAAGCAGCTGCCAGAGACAATAATGCCGTCGCG 6611
786 sThrcIu..ThrSerGluGluGlySerGlyAspLeuAspAsnLeuAspA 802
:|||||:|||||:|||||:|||||:|||||
6612 TTCTGCTAAACATCATCAAGCAGAGCTCCGCG.....6642
802 lallleuGlyAspLeuThrSerSerAspPheTyrAsnAsnSerIleSer 818
```

```
6643 .....AGCAACAGCTGCACAGCACCGAC...TACAGCAGGCCCAATATGC 6681
819 SerAsnGlySerHisLeuGlnGlyThrIleValIlePheGlnGlyThrAs 835
:|||||:|||||:|||||:|||||:|||||
6682 TTCCGAGGAGGATGGCCAGCATGCAGCGAGCTGTGTGGTGGCAGCA 6731
835 nSerLeuGlyLeuLysSerSerGlnSerValGln.....5 847
6732 CAGGGCCCTCCCTTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 6781
847 erIleArgProProTyrAsnArg.....AlaValSerLeuAspSer 860
:|||||:|||||:|||||:|||||:|||||
6782 ACAGCCAAACACCCCGCAGACGCCACCCACCTTCAAGCTCAGCTCA 6831
861 ProValSerValGlySerSerProProValLysAsnIleSerAlaPhe 877
:|||||:|||||:|||||:|||||:|||||
6832 CCCCTCCCAATAGCATGCCCTTACTTCCCGAGGACTCAAGCTCAGCTGC 6881
877 rometLeuProLysGlnProMetLeuGlyLysAsnProArgMetLeuAsp 893
6882 CCTGTGTCCTCA.....6892
894 SerGlnGluAsnTyrGlySerSerMetGlyLysProAsnAlaValTh 910
6893 .....GGGTAAAGCAGCAGCGCCAGGTGACCTCC.....6922
910 rValThrGlnThrProSerSerGlyAspTyrGlyLeuProAsnSerLys 927
6923 .....AACCCCTCCTCA...GACTGCTACACACCCCTTCAAGG 6959
927 lAgIArgMetGluProMetAsnSerAsnSerMetGlyArgProGlyLys 943
6960 CCCCACCTACAGCAGTGAATGGCAATGGCAATGGCAATGGCAAGCAGCG 7009
944 AspTyrAsnThrSerLeuProArgProAlaLeuGlyGlySerIlePro 960
7010 GAGCAGCGCCAGATGGCCAGCTGCAATTTTCAAGGCCAATGCCAAC 7059
960 r.....LeuProLeuArgSerAsnSerIleProGlyAlaArgProVal 975
7060 ACCAATGCCCCCGATGATGATGATGATGATGATGATGATGATGATGAT 7109
975 euGlnGlnGlnGlnGlnMetLeuGlnMetArgProGlyGlyIlePro 991
7110 CCCATGACAGCAGAGTCCAGTGGCATTTGGAGCCAGGATGGCAGCAG 7159
991 tGlyMetGlyAlaAsnProTyrGlyGlnAlaIleAlaSerAsnGlnLeu 1008
7160 AGGATGGCAGCAGACCA.....7178
1008 lYSerTrpProAspGlyMetLeuSerMetGluGlnValSerHisGlyThr 1024
7179 ..CCCTGGAGCCCAAGAGAGATGCTCCAGCCCAACACACACACACAC 7226
1025 GlnAsnArgProLeuLeuArgAsnSerLeuAspAspLeuValGlyPro 1041
7227 ATGCCAAGGCCCATGATGATGATGATGATGATGATGATGATGATGAT 7270
1041 oSerAsnLeu.....1044
7271 TTGTAACATGCTCCCAACACAGATTGGCCAGGTAGTATACGCCAC 7320
1045 .....GluGlyGlnSerAspGluArgAlaLeuLeuAspGlnLeuHisThr 1059
7321 TCAAACACAGCAGCTGTGTCATAAAGCCTTACAAACCTTTTGGGAGCT 7370
1060 LeuLeuSerAsnThrAspAlaThrGlyLeuGluGluIleAspArgAla 1076
7371 CTCAGGTCTCCAGCTCTCCCTGCGAGCAGCAGACGAGCTTAGATCT 7420
1076 u...GlyIleProGluLeuValAsnGlnGlyGlnAlaLeuGluProLys 1092
```



```

1349 .....GlnSerSerGluMetLysGlyTyrProSerGly 1359
      ::::: |||::: :::::
8245 CTGGACTGTAGTAGCGCCAGCAGCCCAACCCCATGCAGAACAAAGGCGATTGTGCC 8294
      :::::

1360 AsnLeuAlaIArgHisSerSerPheSerGlnGlnGlnInpheAlaHISGlnGI 1376
      :|||::: |||::: |||::: |||::: |||:::
8295 AGCCCGGACCAGATTCATGATCTTTCACACTT.....GCTAG 8322
      ::|||::: |||::: |||::: |||::: |||:::
8333 CATTCGA.....GGCATGGCAAAACCTCCACATGGTGCAAGC 8366
      ::|||::: |||::: |||::: |||::: |||:::

seq_name: /cgn2_6/plodata/1/ina/5B_COMP.seq:US-08-785-310A-4

seq_documentation_block:
Sequence 4, Application US/08785310A
Patent No. 5840532
GENERAL INFORMATION:
APPLICANT: McKnight, Steven L.
APPLICANT: Russell, David W.
TITLE OF INVENTION: Neuronal PAS Domain Protein
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/785,310A
FILING DATE: 21-JAN-1997
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UTSD:1226
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO.: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 4184 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-785-310A-4

alignment_scores:
Quality: 264.00 Length: 1017
Ratio: 0.532 Gaps: 53
Percent Similarity: 48.771 Percent Identity: 21.141

alignment_block:
US-09-041-994-2 x US-08-785-310A-4 ..

Align seg 1/1 to: US-08-785-310A-4 from: 1 to: 4184

34 GIUUYARGARARGeluglncJuserySYRIleGlucIuealagl 50
|||||:::|||||::: |||::: |||::: |||::: |||:::
624 GAGAAGAAGCGTGGCGACGATTCATGTTCTCATCAAAGAGCTCACGTC 673
|||||::: |||::: |||::: |||::: |||::: |||:::
50 uleuIeSerAlaAsnIeuSerSeriIeaSpaInpheASnVallysProA 67
:::|||||::: |||::: |||::: |||::: |||::: |||:::
674 CAGGCTCCCTGTGTAACACTGG.....AAATGG 702

```

[illegible]



```

2937 .....GGAACCTGGGCTATGCA 2954
906 ASnArGsnVnAlThrVnAlThrGlnThrProSerSerGlyAspTrpGlyLe 922
    ::::|::|::|::|::|
2955 GCCACACAGCTTACACCTCCGCAACCCCGCAGC..... 2987
922 uPrAnSerLysAlaGlyArGmetGluProMetAnSerAnSerMetG 939
    ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
2988 .CCCTCCCGCAGGCTCAGCCGGCTGCA.....GAGTCTTAAGTCTCC 3030
939 lYArProGlyGlyAspTrpLysThrSerLeuProArGProAlaLeuGly 955
    ::::|::|::|::|::|::|::|::|::|::|
3031 AGCAGCATCAGGCTATATC.....CCTACACTGGA 3062
956 GlySerIleProThrIleuProLeuArGSerAnSerIleProGly 970
    ::::|::|::|::|::|::|::|::|::|::|
3063 .....GTCCAGACGCACACACAGCTTAAACCAATGAAACGGGGT 3101

seq_name: /cgn2_6/ptodeta/1/lna/5B_COMB.seq:US-08-480-473B-1

seq_documentation_block:
: Sequence 1, Application US/08480473B
: Patent No. 5882914
: GENERAL INFORMATION:
: APPLICANT: Semenza, Gregg L.
: TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
: NUMBER OF SEQUENCES: 64
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 4225 Executive Square, Suite 1400
: CITY: La Jolla
: STATE: CA
: COUNTRY: USA
: ZIP: 92037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/480,473B
: FILING DATE: 06-JUN-1995
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Halle, Lisa A.
: REGISTRATION NUMBER: 38,347
: REFERENCE/DOCKET NUMBER: 07265/053001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619/678-5070
: TELEFAX: 619/678-5099
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3736 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA
: US-08-480-473B-1

alignment_scores:
Quality: 258.00 Length: 915
Ratio: 0.604 Gaps: 41
Percent Similarity: 46.667 Percent Identity: 19.781

alignment_block:
US-09-041-994-2 x US-08-480-473B-1 ..
Align seg 1/1 to: US-08-480-473B-1 from: 1 to: 3736

1 MetcErGlyLeuGlyGluAnLeuAsp.....ProLeuAlaSerAspse 15
29 ATGAGGAGCGCGCGCGCGGAACGACGAAAGAAAGATAGTTCTGCAAG 78

```

```

2937 .....GGAACCTGGGCTATGCA 2954
906 ASnArGsnVnAlThrVnAlThrGlnThrProSerSerGlyAspTrpGlyLe 922
    ::::|::|::|::|::|
2955 GCCACACAGCTTACACCTCCGCAACCCCGCAGC..... 2987
922 uPrAnSerLysAlaGlyArGmetGluProMetAnSerAnSerMetG 939
    ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
2988 .CCCTCCCGCAGGCTCAGCCGGCTGCA.....GAGTCTTAAGTCTCC 3030
939 lYArProGlyGlyAspTrpLysThrSerLeuProArGProAlaLeuGly 955
    ::::|::|::|::|::|::|::|::|::|::|
3031 AGCAGCATCATGGGTATATC.....CCTACACTGTGA 3062
956 GlySerIleProThrIleuProLeuArGSerAnSerIleProGly 970
    ::::|::|::|::|::|::|::|::|::|::|
3063 .....GTCCAGACGCACACACGACTTATACCAATGAAACGGGGT 3101

seq_name: /cgn2_6/ptodeta/1/lna/5B_COMB.seq:US-08-480-473B-1

seq_documentation_block:
: Sequence 1, Application US/08480473B
: Patent No. 5882914
: GENERAL INFORMATION:
: APPLICANT: Semenza, Gregg L.
: TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
: NUMBER OF SEQUENCES: 64
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 4225 Executive Square, Suite 1400
: CITY: La Jolla
: STATE: CA
: COUNTRY: USA
: ZIP: 92037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/480,473B
: FILING DATE: 06-JUN-1995
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Halle, Lisa A.
: REGISTRATION NUMBER: 38,347
: REFERENCE/DOCKET NUMBER: 07265/053001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619/678-5070
: TELEFAX: 619/678-5099
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3736 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA
: US-08-480-473B-1

alignment_scores:
Quality: 258.00 Length: 915
Ratio: 0.604 Gaps: 41
Percent Similarity: 46.667 Percent Identity: 19.781

alignment_block:
US-09-041-994-2 x US-08-480-473B-1 ..
Align seg 1/1 to: US-08-480-473B-1 from: 1 to: 3736

1 MetcErGlyLeuGlyGluAnLeuAsp.....ProLeuAlaSerAspse 15
29 ATGAGGCGCGCGCGCGGCAAGCAAGAAAGATAGATTCTGCAAG 78

```

```

15 rArgLysArgLysLeuProCysAspThrProGlnGlnGlyLeuThrCys 32
   |||||
79 TCGAAAGAAAG.....TCTCAG 98

32 eArgGlyLysArgArgArgGlnGlnGlnSerLysTyrIleGlnGln 48
   :|||
99 ATGCAGCCGATCTCGCGAGTAAGAAATCGAAGTTTATATACGCTT 148

49 AlaGlnLeuIleSerAlaAsnLeuSerAspIleAspAsnPheAsnVal 65
   |||||
149 GCTCAT.....CAGTTGCCACTCCACATATAGTAGTTCCGA 186

65 sProAspLysCysAlaIleLeuLysGluThrValArg.....GlnIle 80
   |||||
187 TCTGTATAGGCTCTGTGATGAGCTTACCATCTATTGCGGTGA 236

80 rGlnIleLeuGlnGlnGlnLysThrIleSerAsnAspAspValGln 96
   |||||
237 GGAACCTCTGTGATGCTGT...GATTGGATATTGAAGATGACATG... 280

97 LysAlaAspValSerSerThrGlnGlnValIleAspLysAspSer 113
   |||||
281 .....AAGCACAGAT 291

113 uGlyProLeuLeuGlnAlaLeuAspGlyPheLeuPheValAlaAsn 130
   :|||
292 GAATGCTTTTATTTGAAGCCTGTGATGTTTGTATGTTCTCACAG 341

130 rGlnAlaAsnIleValPheValSerGluAsnValThrGlnTyrLeuGln 146
   :|||
342 ATGATGCGTACATGATTACATTCTGATATGTAACAATAATACATGGA 391

147 TyrLysGlnGlnAspLeuValAsnThrSerValTyrAsnIleLeuHisG 163
   |||||
392 TTAACTGATTGAACCTGACACACAGCTGTTGATTACTACATCC 441

163 uGlnAspArgLysAspPheLeuLysAsnLeuProLysSerThrValAsnG 180
   |||||
442 ATGTACCACTGAGAAATGAGAAATCTTACACACAGA.....ATG 485

180 LysValSerThrAsnGlnProGlnArgGlnLysSer..... 192
   :|||
486 GCGTTGTGAAAAGGTAAGAACAACACACAGCAAGCTTTTTC 535

193 .....HisThrPheAsnCys.... 197
   :|||
536 AGAATGAAGTGTACCTTAAGTACGAGAGAAAGTATGACATATAAGTC 585

198 .....ArgMetLeuMetLysThrProHisAspIleLeuGluAspI 211
   :|||
586 TGCACATGGAAGGTATGTGACTGCACAGGCCACATTCACGTATATGATA 635

211 LeuAsnAlaSerProGlnMetArgGlnArgTyrGlnThrMetGlnCysP 227
   |||||
636 CCAACAGTAACCAACT.....CAGTGTGG 661

228 AlaLeuSerGlnProArgAlaMetMetGlnGlnGluAspLeuGlnSe 244
   :|||
662 TATAGAAACCACT.....ATGAC 681

244 rCysMetIleCysValAlaArgArgIleThrThrGlnGluArgThrPhe 261
   :|||
682 CTGCTGTGTGCTGATTGTGAAACCACT.....CCTCAC 716

261 rSerAsnProGln.....SerPheIleThrArgHis 271
   |||||
717 CATCAAAATTTGAATTCCTTAGATACGACAGACTTCTCTCAGTGCACAC 766

272 AspLeuSerGlyLysValValAsnIleAspThrAsnSerLeuArgSer 288
   :|||
767 AGCCTGATATGAATTTCTTATTTGTATGA.....AGAATTAC 807

```

```

288 rMetArgProGlyPhe.....GluAspIleIleArgCysIleGln 303
   :|||
808 CGAATTGATGGATATGAGCCAGAAAGAACTTTAGCCGCTCAATTTATG 857

303 rGpPhePheSerLeuAsnAspGlyInsSerTrpSerGlnLysArgHisTyr 319
   :|||
858 AATATTATCATGCTTTGGACCTGTGATCATCTGACCAAAATCATAT... 904

320 GlnGluAlaTyrLeuAsnGlyHisAlaGlnThrProValTyrArgPhe 336
   :|||
905 ...GATATGTTACTAAAGACAACTCAACACAGACAGATACAGATGCT 951

336 rLeuAlaAspGlyThrIleValThrAlaGlnThrLysSerLysLeuPhe 353
   |||||
952 TGCCAAAGAGGTGATATGCTGTGGTTGAACCTCAACCACTGATAT 1001

353 rGAsnProValThrAsnAspArgHisGlyPheValSerThrHisPheLeu 369
   |||||
1002 ATACACCAAGAAATTCACACAGCTGATGTGTGTAATTACGTT 1051

370 GlnArgGlnGlnAsnGlyTyrArgProAsnProAsnProValGlnGln 386
1051 ..... 1051

386 rIleArgProProMetAlaGlyCysAsnSerSerValGlyGlyMet... 401
   |||||
1052 .....GTGAGGTATTTATTC 1067

402 .....SerMetSerPro 405

1068 AGCAGACTGATTTTCCTCCCTCAACAACAGATGTGCTTAAACG 1117

406 AsnGlnGlyLeuGlnMetProSerSerArgAlaTyrGlyLeuAlaAsp 422
   :|||
1118 GTTGATCTTCAGATATGAAATGACTCAGCATTCACCAAGTTGAATC 1167

422 oSerThrThrGlnGlnMetSerGlyAlaArgTyrGlyLysSerAsnI 439
   :|||
1168 AGAAGATACAGTACCTCTTGGACAACTTAAGAGAACTGATGCTT 1217

439 LeuAlaSerLeuThrProGlyProGlyMetGlnSerProSerSerTyrGln 455
   :|||
1218 TAACTTGTGCGCCCGCCAGCGCTGA.....GACCAATCATTA 1255

456 AsnAsnAsnTyrGlyLeuAsnMetSerSer..... 465
   :|||
1256 TCTTTAGATTGTCGACAGACACAGAAACTGATGACCAAGCACTTGA 1305

466 .....ProProHisGlySerProGlyLeu...AlaProAsnGlnGln 478
   |||||
1306 GGAAGTACCATTTATATATGATGTATGCTCCCTCCACCAAGCAAAAT 1355

479 .....AsnIleMetIleSerProArgAsnArgGlySerProLys 491
   |||||
1356 TACAGATATATAATTTGGCAATGCTCCA..... 1384

492 IleAlaSerHisGlnPheSerProValAlaGlyValHisSerProMetAl 508
   :|||
1385 .....TTAACCCACCGCTGAACGCCAACGCCACTTCG 1416

508 sSerSerGlyAsnThrGlyAsnHisSerPheSerSerSerLeuSera 525
   |||||
1417 AAGTAGTCTGAC..... 1429

525 AlaLeuGlnAlaIleSerGlnGlyValGlyThrSerLeuLeuSerThrLeu 541
   |||||
1430 .....CTGCACTCAATCAAGAAAGTTGCATTAATTA..... 1462

542 SerSerProGlyProLysLeuAsnAsnSerProAsnMetAsnIleThrGln 558
   :|||
1463 ...GAACCAAAATCCAGAG.....TCACGTGAACCTTCTTTTACCAT 1500

558 nPro.....SerLysV 562

```

```

1501 GCCCGAGATTCAGATCAGACACCTAGTCTCCATGAGACACCTAGAC 1550
      |||
562 aSerAsnGlnAspSerIysSerPro.....LeuGlyPheTyrCysAsp 576
      |||::: |||
1551 AAAGTTCACCTGAGCCTATAGTCCAGAGAAATATTGTTTATGAGAT 1600
      |||::: |||
577 GlnAsnProValGluSerSerMetCysGlnSerAsnSerArgAspHisLe 593
      |||::: |||
1601 AGTGATATGGTCATGAATTCAGATTGGAATTGGTAAAGAACTTTTTC 1650
      |||::: |||
593 uSerAspIysGluSerIysGluSerSerValGluGlyAlaGlnGlnA 610
      |||::: |||
1651 TGAAGACACAGAAAGCAAG..... 1669
610 rGlyProLeuGluSerIysGlyHisIysIysLeuLeuGlnLeuLeuThr 626
      |||::: |||
1670 ..AACCCATTTCCTAGTCCAGACAGATTTAGACTTGGAGATGTTAGCT 1717
      |||::: |||
627 .....CysSerSerAspAsp.....ArgGlyHisSerSerIle 637
      |||::: |||
1718 CCTATATCCCAATGATGATGACTTCAGTTACGTTCTCTGATCAGTT 1767
      |||::: |||
637 uThrAsnSerProLeuAspSerSer.....CysL 647
      |||::: |||
1768 G.....TCCACCTTAGAAAGCAGTTCCGCAAGCCCTGAAGCCGACATC 1811
647 yGluSerSerValSerValIThrSerProSerGlyValSerSerSerThr 663
      |||::: |||
1812 CTCAAGACACAGCTTACAGTATTCACGACAGACCTCAATATACAACTT 1861
      |||::: |||
664 SerGlyGlyValSerSerThrSerAsnMetHisGlySerLeuLeuGln 680
      |||::: |||
1862 GCTAATGCCACCACTACCCACTGCTGATGATTAATAAAGCTGAC 1911
680 uIysHisArgIleLeuHis...LysLeuLeuGlnAsnGlyAsnSerPro 696
      |||::: |||
1912 AAAAGCCGCTATGGAAGACATTAATATGATGATGATTCATCTCCCTTA 1961
696 lAgIuValAlaIysIleThrAlaGlnAlaThrGly.....LysAsp 709
      |||::: |||
1962 CCCACATATACATAAGAACTACTAGTCCACATCATCACCATATAGAGAT 2011
710 ThrSerSerIleThrSer.....CysGlyAspGly..... 719
      |||::: |||
2012 ACTCAAGTCGAGACGCTCACCACAAACAGACGAAAGAGTCTATAGA 2061
719 ..... 719
2065 ACAGACAGAAAAATCTCATCCAAAGCCCTAACGCTTATCTGTCGCTT 2111
720 .....AsnValValIysGlnGlnGlnLeuSerProLys..... 730
      |||::: |||
2112 TGAGTCAAAAGAACTACAGTTCCTGAGAGAACTAAATCCAAAGATACTA 2161
731 .....LysIysGlnAsnAsnAlaLeuLe 738
      |||::: |||
2162 GCTTTCAGAAATGCTCAGAAAGCGAAATGGAACATGATGATGCTTCACT 2211
738 u.....ArgTyrLeuLeuAspArgAspAspProSer 749
      |||::: |||
2212 TTTTCACAGCAGTAGAATGGAACATTTATTACAGCAGCCAGACGATCARG 2261
749 sPAlaLeuSerIysGluLeuGlnProGlnValGluGlyVal..... 762
      |||::: |||
2262 CAGCTACTACTACTTCTTGAAACGGTAAAGGATGCAAAATCTACT 2311
763 ...AspAsnIysMetSerGlnCysThrSerSerThrIleProSer 776
      |||::: |||
2312 GAACAGAAATGGAATGAGCAAAAGACAAATATTATTATATCCCTCT 2356
      |||::: |||
seq_name: /cgn2_6/plodata/1/lna/6a_COMB.seq:US-08-915-213-1

```

```

seq_documentation_block:
; Sequence 1, Application US/08915213
; Patent No. 6020462
; GENERAL INFORMATION:
; APPLICANT: Semenza, Gregg L.
; TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,213
; FILING DATE: 20-AUG-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/480,473
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/053001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3736 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-915-213-1

alignment_scores:
Quality: 258.00 Length: 915
Ratio: 0.604 Gaps: 41
Percent Similarity: 46.667 Percent Identity: 19.781

alignment_block:
US-09-041-994-2 x US-08-915-213-1 ..
Align seg 1/1 to: US-08-915-213-1 from: 1 to: 3736
1 MetSerGlyLeuGlnGlyGlnAsnLeuAsp.....ProLeuAlaSerAsp 15
|||::: |||
29 ATGGAAGGCGCGCGCGCGCGGACGACGACGAAAGAAATATGATTCGAAG 78
|||::: |||
15 rArgIysArgIysLeuProCysAspThrProGlyGlnGlyLeuThrCys 32
|||::: |||
79 TCGAAAGCAAGAAAG.....TCTCGAG 98
|||::: |||
32 erGlyGluIysArgArgArgGlnGlnGlnIysSerIleGlnLeuLeu 48
|||::: |||
99 ATGCAAGCCAGATCTCGCGAGAGTAAAGATTCGAATTTTATATGAGCTT 148
|||::: |||
49 AlagIuLeuIleSerAlaAsnLeuSerAspIleAspAsnPheAsnVal 65
|||::: |||
149 GCTCAT.....CAGTTGCCACTTCCACATATATGATGATTCGCA 186
|||::: |||
65 sProAspIysCysAlaIleLeuIysGluThrValArg.....GlnIleA 80
|||::: |||
187 TCTGTATAGGCGCTCTGTCATGAGGCTTACCATCATATTTGCTGTGCA 236
|||::: |||

```





```

228 AlaLeuSerGlnProArgAlaMetMetGluGluGlyGluAspLeuGln 244
    ::::|::|
662 TATAAGAAACCACT.....ATGAC 681
244 rcysmetIleCysValAlaArgArgIleThrThrGlyGluArgThrPhe 261
    :|::|::|::|::|::|
682 CTGCTTGGTGGTGAATTTGAAACCACT.....CCTCACC 716
261 roSerAsnProGlu.....SerPheIleThrArgHis 271
    |||||::|::|::|::|::|
717 CATCAAAATTTGAAATTCCTTAGATAGCAAGACTTCTCTCAGTCAGCAC 766
272 AspLeuSerGlyLysValValAsnIleAspThrAsnSerLeuArgSer 288
    ::::|::|::|::|
767 ACCCTGGATATGAAATTTCTTATTGTGTGAA.....AGAAATTAC 807
288 rMetArgProGlyPhe.....GluAspIleIleArgArgCysIleGln 303
    :|::|::|::|::|::|::|
808 CGAAATTGATGGATATGAGCCAGAAAGAACTTTAGCGCGCTCAATTTATG 857
303 rgPheSerLeuAsnAspGlyGlnSerTrpSerGlnLysArgHisTyr 319
    ::::|::|::|::|::|::|
858 AATATTATATGCTTTGGACTGTGATCATGTACCAAAACTCATCAT... 904
320 GlnGluAlaTyrLeuAsnGlnHisAlaGluThrProValTyrArgPhe 336
    ::::|::|::|::|::|::|
905 ...GATATGTTTACTAAAGACAGACACAGACAGACAGTACAGATGCT 951
336 rLeuAlaAspGlyThrIleValThrAlaGlnIleLysSerLysLeuPhe 353
    |||||::|::|::|::|::|
952 TGCCAAAAGAGGTGGATATGCTGGGTGAACCTCAAGCACTGTCATAT 1001
353 rgAsnProValThrAsnAspArgHisGlyPheValSerThrHisPheLeu 369
    |||::|::|::|::|::|
1002 ATAAACACCAAGATTCTCAACCAAGTCATGTGATGTGAATTACGTT 1051
370 GlnArgGluGlnAsnGlyTyrArgProAsnProAsnProValGlyGln 386
1051 ..... 1051
386 ylleArgProPromeAlaGlyCysAsnSerSerValGlyGlyMet.... 401
    ||::|::|::|::|
1052 .....GTGAGTGGTATTATATC 1067
402 .....SerMetSerPro 405
1068 AGCAGCACTTGATTTTCTCCTTCAACAAACAGAAATGTCTCTTAACCG 1117
406 AsnGlnGlyLeuGlnMetProSerSerArgAlaTyrGlyLeuAlaAsp 422
    ::::|::|::|::|
1118 GTTGAATCTTCAGATATGAAATGACTCAGTATTCACCAAGTTGAATC 1167
422 oSerThrThrGlyGlnMetSerGlyAlaArgTyrGlyGlySerSerAsn 439
    ::::|::|::|::|::|
1168 AAGAAGTAAAGTAGCCTCTTGACAAACTTAAGAAGAACTGATGCTT 1217
439 leaIleSerLeuThrProGlyProGlyMetGlnSerProSerSerTyrGln 455
    ::::|::|::|::|::|
1218 TACCTTTCCTGGCCCGCCAGCGCTGCA.....GACAAATCATATA 1255
456 AsnAsnAsnTyrGlyLeuAsnMetSerSer..... 465
    ::::|::|::|::|
1256 TCTTTAGATTTTGGCAGACAGACAGAAACTGATGACCAGCAACTTGA 1305
466 .....ProProHisGlySerProGlyLeu...AlaProAsnGlnGln 478
    |||::|::|::|::|
1306 GGAAGTACCATATATATGATGATATGCTCCCTCACCACCAAGAAAT 1355
479 .....AsnIleMetIleSerProArgAsnArgGlySerProLys 491
    ||::|::|::|::|
1356 TACAGAAATATAATTTGGCAATGTCTCA..... 1384

```

```

492 IleAlaSerHisGlnPheSerProValAlaGlyAlaHisSerProMetAl 508
    ||::|::|::|::|
1385 .....TTACCCACCGCTGAAGCCCAAGCCAAACCACTTGC 1416
508 aSerSerGlyAsnThrGlnLysHisSerPheSerSerSerLeuSerA 525
    |||||::|::|
1417 AAGTATGCTGTAC..... 1429
525 leaLeuGlnAlaIleSerGluGlyValGlyThrSerLeuLeuSerThrLeu 541
    ||::|::|::|::|::|
1430 ....CCTCAGCTCAATCAAGAAAGTTGCATTAAATTA..... 1462
542 SerSerProGlyProLysLeuAspAsnSerProAsnMetAsnIleThrG 558
    ::::|::|::|::|::|
1463 ...GAACCAATCCAGAG.....TCACGTGAACCTTCTTTTACCAT 1500
558 nPro.....SerLysVal 562
    :|::|
1501 GCCCCAGATTCAAGATCAGACACCTAGTCTTCCGATGGAGCACTAGAC 1550
562 aIleAsnGlnAspSerLysSerPro.....LeuGlyPheTyrCysAsp 576
    ||::|::|::|::|
1551 AAGATTACCTGAGCGCTAATAGTCCAGTGAATATTGTTTATGTGAT 1600
577 GlnAsnProValGlnSerSerMetCysGlnSerAsnSerArgAspHisLe 593
    ::::|::|::|::|
1601 AGTGATATGGTCAATGAATTCAGTTGGAAATGTGTAAGAACTTTTGC 1650
593 uSerAspLysGluSerLysGluSerSerValGluGlyAlaGluAsnGln 610
    ::::|::|
1651 TGAAGACACAGAAACCAAG..... 1669
610 rgGlyProLeuGluSerLysGlnHisLysLysLeuLeuGlnIleLeuThr 626
    ::::|::|::|::|
1670 ..AACCAATTTCTACTCAGACAGATTTAGACTGTGAGATGTAGCT 1717
627 .....CysSerSerAspAsp.....ArgGlyHisSerSerLe 637
    |||||
1718 CCTATATCCCAATGATGATGACTTCCAGTTACGTTCTTCCGTGATCAGT 1767
637 uThrAsnSerProLeuAspSerSer.....CysL 647
    | |||||::|::|
1768 G.....TCACCATTAAGAAAGCAGTCCCGCAAGCCCTGAAGCGCAAGTC 1811
647 ysgLysSerSerValSerValThrSerProSerGlyValSerSerSerThr 663
    ::::|::|::|::|::|
1812 CTCAAAGCACAGTTACAGTATTCAGACAGACTCAATACAAAGAACTTACT 1861
664 SerGlyGlyValSerSerThrSerAsnMetHisGlySerLeuLeuGln 680
    ::::|::|::|::|
1862 GCTAATGCCACCACTACACCTGCCACCTGATGAATTAATAACAGTCAC 1911
680 uLysHisArgIleLeuHis...LysLeuLeuGlnAsnGlnLysAsnSerPro 696
    ||| ||::|::|::|
1912 AAAAGACCGTATGGAAGACATTAATAATATGATTCGATCTCATCTCCGA 1961
696 leaGluValAlaLysIleThrAlaGlnAlaThrGly.....LysAsp 709
    ::::|::|::|::|::|
1962 CCCACATACATTAAGAAACTACTAGTCCACATCATCAACATATAGAGAT 2011
710 ThrSerSerIleThrSer.....CysGlyAspArgLys..... 719
    ||::|::|::|
2012 ACTCAAGTGGACAGCCCTCACCAAAACAGAGCAAGAAAGAGTATATGA 2061
719 ..... 719
2062 ACAGACAGAAATATCATCAAGAAGCCCTAAGCTGTTATCTGTGCGTT 2111
720 .....AsnValValLysGlnGluGlnLeuSerProLys..... 730
    ::::|::|::|::|
2112 TGAGTCAAGAACTACAGTCTCTGAGAGAAAGAACTAAATCCAAAGATACTA 2161
731 .....LysLysGluAsnAsnAlaLeuLe 738

```



```
2162 GCTTGCAGATGCTCAGAGAAAGCAAAATGAAACATGATGTTCACT 2211
738 u.....ArgTyleuLeuaspargaspProsera 749
2212 TTTTCAGACAGTAGAATTTGAAACATTATTACAGCCAGCAGATCARG 2261
749 sPAlaleuSerlysgluLeuGlnProGlnValGlnGlyVal..... 762
2262 CAGCTACTACTACATCTTCTTGAAACGGTAAAGATGCAAAATCTACT 2311
763 ...AspAsnlyMetSerGlnCysThrSerSerThrIleProSer 776
2312 GAACAGAAATGAAATGAGCAAAAGCAATATTATTATACCTCT 2356
seq_name: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq: PCT-US96-10251-1

seq_documentation_block:
: Sequence 1, Application PC/TUS9610251
: GENERAL INFORMATION:
: APPLICANT: The Johns Hopkins University School of Medicine
: TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
: NUMBER OF SEQUENCES: 35
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 4225 Executive Square, Suite 1400
: CITY: La Jolla
: STATE: CA
: COUNTRY: USA
: ZIP: 92037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US96/10251
: FILING DATE: 06-JUN-1996
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Halle, Lisa A.
: REGISTRATION NUMBER: 38,347
: REFERENCE/DOCKET NUMBER: 07265/053W01
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619/678-5070
: TELEFAX: 619/678-5099
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3736 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA
: PCT-US96-10251-1

alignment_scores:
  Quality: 258.00      Length: 915
  Ratio: 0.604        Gaps: 41
  Percent Similarity: 46.667      Percent identity: 19.781

alignment_block:
US-09-041-994-2 x PCT-US96-10251-1 ..
Align seg 1/1 to: PCT-US96-10251-1 from: 1 to: 3736

1 MetSerGlyLeuGlnLeuAsp.....ProLeuAlaSerAspse 15
|||||
29 ATGAGAGGCGCGCGCGCGAGACGACAAAGAAAGATTAAGTTCTGAACG 78
15 rArglysrArglyLeuProCysAspThrProGlyGlnGlyLeuThrCys 32
|||||
79 TCGAAAGAGAAAG.....TCGCGAG 98
```

```
32 erGlyIuIysrArgArgGluGlnGlnSerIlystrIleGluLeu 48
:::::
99 ATGCACCCGAGATCTCGCGAGAGTAAGATCTGAAGTTTATGAGCTT 148
49 AlaGluLeuIleSerAlaAsnLeuSerAspIleAspAsnPhenAlly 65
|||||
149 GCTCAT.....CAGTTGCCACTTCCACATTAATGATGCTTCGA 186
65 sProAspLysCysAlaIleLeuLysGluThrValArg.....GlnIleA 80
187 TCTTGATTAAGGCTCTGTGATGAGGCTTACATCAGCTATTTCGGTGTA 236
80 rGlnIleLysGluGlnGlyIlystrIleSerAsnAspAspAlaGln 96
|||||
237 GGAAACTTCTGATGCTGCT...GATTGGAATATTCAGATCAGATG... 280
97 LysAlaAspValSerSerThrGlyGlnGlyValIleAspLysAspSerIle 113
|||||
281 .....AAAGCACAGAT 291
113 uGlyProLeuLeuGlnAlaLeuAspGlyPheLeuPheValAlaAsnA 130
:::::
292 GAATTCCTTTTATTTCGAAGCCTTGATGCTTTGTTATGCTTCACAG 341
130 rGluAlaAsnIleValIlePheValSerGluAsnValThrGlnIlyLeuGln 146
:::::
342 ATGATGGTGACATGATTTTACATTTCTGATTAAGTGTAACAAATACAGGA 391
147 TyrLysGlnGluAspLeuValAsnThrSerValTyrAsnIleLeuHisG 163
|||||
392 TTAACCTCACTTGAACCTAGCTGACACAGCTGTGTTGATTTTACATCC 441
163 uGluAspArgLysAspPheLeuLysAsnLeuProLysSerThrValAsnG 180
|||||
442 ATGTGACCATGAGGAATGAGAAATGCTTACACACAGA.....AATG 485
180 LysValSerThrAsnGluProGlnArgGlnLysSer..... 192
|||||
486 GCCTTGTAAGAAAGGCTAAAGAACCAACACAGCAGGAAGCTTTTTCYC 535
193 .....HisThrPheAsnCys..... 197
536 AGAATGAAGTACCTTACCTAGCCGAGAGAGAACTATTAACATTAAGTC 585
198 .....ArgMetLeuMetLysThrProHisAspIleLeuGluAspI 211
:::::
586 TGCACATGGAAGATATTCACCTGCACAGCCACCATTCACGTATATGATA 635
211 IeAsnAlaSerProGluMetArgGlnArgTyrGluThrMetLincysPhe 227
|||||
636 CCAACAGTAACCAACCT.....CAGTGTGGG 661
228 AlaLeuSerGlnProArgAlaMetMetGluGluGlyLysPheLeuGln 244
:::::
662 TATTAAGAAACCACT.....ATGAC 681
244 rCysMetIleCysValAlaArgArgIleThrThrGlyGluArgThrPhe 261
|||||
682 CAGCTTGCTGCTGATTTTGTGAACCATT.....CCTCAC 716
261 rSerAsnProGlu.....SerPheIleThrArgHis 271
|||||
717 CATCAAAATATGAAATTCCTTAGATAGCAGAACTTCTCAGCTGACAGC 766
272 AspLeuSerGlyLysValAlaAsnIleAspThrAsnSerLeuArgSer 288
:::::
767 ACCTTGATATCAAAATTTCTTATTTGTGATGAA.....AGAAATAC 807
288 rMetArgProGlyPhe.....GluAspIleIleArgArgCysIleGlnA 303
|||||
808 CGAATTGATGGATATAGCCAGAGAAAGAACTTTAGCCGCTCAATTTATG 857
```



```

857 SerLeuAspSerProValSerValGlySerSerProValIAsnIL 873
      ||||| |||||:::||||::: |||||:::|||||:::
5644 AGTTGGCTTCTCTACTCTAGCACACCCGGGACTCTTACACACGAC. 5691a
      ||||| |||||:::||||::: |||||:::|||||:::
873 eSerAlaPheProMetLeuProLysGlnProMetLeuGlyLysProA 890
      ||||| |||||:::||||::: |||||:::|||||:::
5692 .....CCGACACACCCCAACACCCACACCCCGCCAGCCTC 5731b
      ||||| |||||:::||||::: |||||:::|||||:::
890 rgMetMetAspSerGlnGluAsnTyrLysSerMetGlyLysProAsn 906
      ::|||::: |||||:::|||||:::
5732 AG.....CCTTCA 5739c
      ||||| |||||:::||||::: |||||:::|||||:::
907 ArgAsnValThrValThrGlnThrProSerSerGlyAspTrpGlyLeu 923
      |||||:::|||||::: |||||:::|||||:::
5740 .....CCTGTAAACATGTCACACGAA.....GGCTTCCC 5768d
      ||||| |||||:::||||::: |||||:::|||||:::
923 oAsnSerLysAlaGlyArgMetGluProMetAsnSerAsnSerMetGlyA 940
      ||||| |||||:::||||::: |||||:::|||||:::
5769 TAAT.....GTAGCCCGGAGTACCGCCCAACAAAGATAGTGTCTCTGGGA 5812e
      ||||| |||||:::||||::: |||||:::|||||:::

```

940 *rgPr1yglYasphyrsntThrsLeuProaPgPr1aLeuGlYgl* 956  
5813 AGCT.....ACCAACAGGGCCAGCTCCCCACCCCTGCC 5850  
957 *serLeProThLeuProLeuArSerSerLeProGlYalaArp* 973  
5851 CAGCCCCACCTGCA.....GCATGAGAGAGCGCGGCA 5885  
973 *oVal*.....LeuGlnGlnGlnGlnMetLeuGln..... 983  
5886 AATTGAACGTGAGGCCACAGACACACACCTATTACGAGCAAACTCA 5935  
984 .....MetArpProGluGluLeuPrometGly..... 992  
5936 ACAATGGCATGCCCCAGACCTGACGGTATGGGGACCCAGAACCA 5985  
993 .....MetGlyAlaAsn.....ProTyrGly 1000  
5986 ATGACCTCCTGGGCGCTGAATGTGCCCGTCCCAACCAAGTCATGGGCC 6035  
1000 *nAlaAlaAlaSerAsnGlnLeuGlySerTrpProAspGlyMetLeuSer* 1017  
6036 TGTCAATGTATTATGACCACTGGCGCATGGACAGACCACTATCCCTC 6085  
1017 *etGlnGlnValSerHisGlyThrGlnAsnArpProLeuLeuAlaGlnSer* 1033  
6086 AGCAGACCGCATGTCAGACG...ATGCCAAGCCCTGTAAATTCATCATG 6132  
1034 *LeuAspAspLeuValGlyProProSerAsnLeuGlnGlnSerAspGly* 1050  
6133 GCCCAGGACAGCTGGCTGGCCACAGGATGCCCAATGTGACGCCAAACG 6182  
1050 *u*.....ArgAlaLeuLeuAspGlnLeuHisThrLeuLeuSerA 1063  
6183 GAGCATCTGCCCAAGGCCCTGACAGACCTGTACGGACCTTAATGCTAC 6232  
1063 *snThrAspAlaThrGlyLeuGlnGluIleAspArgAlaLeu*..GlyTle 1078  
6233 CCAGCTCTCTCAGACGACGACGAGGTGCTGCAATCTTAAATCAAC 6282  
1079 *ProGluLeuValAsnGlnGlyAlaLeuGluProlYsGlnAspAlaPh* 1095  
6283 CCACAGCTAATG.....GCAGCTTT 6302  
1095 *etGlnGlnGlnAlaAlaValMetMetAspGlnYsAlaGlyLeuTyrG* 1112  
6303 CATCAACAGCGCACAGCCCAAGTATGTGGCAATCAGCGCTGCATGACG 6352  
1112 *ly*.....GlnThrTyrProAlaGlnGlyProPrometGlnGly 1124  
1125 *GlyPheHisLeuGlnGlyGlnSerProSerPheAsnSerMetMetAsnGly* 1141  
6397 GGGATGACACCAAGCTGATGTTGGCAAAACCGAAGCATGCAAGCTGG 6446  
1141 *nMetAsnGlnGlnYAsnPheProLeuGln*.....GlyMetH 1154  
6447 TGTGCCAGCGCGGTGGTCCCTCCACCAACCAACGACATGGGAGCGCTGA 6496  
1154 *isProArgAla*.....AsnIleMetArgProArgThrAsn..... 1165  
6497 ATCCCCAGGAGCAAGCTGTGAACATCAATGAACCCAGGACACACCCAC 6546  
1166 .....ThrProGlnGlnLeuArgMetGlnLeuGlnGlnArgTle 1178  
6547 ATGACAAACATGAATTCACAGTACCGAATAATGTGAGGACACAGCTCT 6596  
1178 *uGlnGlyGlnGlnPheLeuAsnGlnSerArgGlnAlaLeuGluLeuTyrSM* 1195  
6597 ACAGCAACCAAGACAGACAGACCAACACAGACAGACAGACAGACAAAC 6646  
1195 *etGluAsnProThr*...AlaGlyGlyAlaAlaValMetArgPrometMet 1210

```

6647 AATAATGTCACAGCTTGGCCGGGGGSCATGGCGGACACAGCCAGTTCACG 6696
1211 GlnProGln.....GlnGlyPheLeuAsnAlaGlnMetValAlaGln 1225
6697 CAGCCCAAGAGACCTGGAGGTATGCCCCAGCATCGACGACGAGCAAGCAT 6746
1225 gSerArgGluLeuLeuSerHisHisPheArgGlnGlnArgValAlaMet 1242
6747 GCAMAGCACCTCCCATCCAGGCGAGCTCCATGGGCGCAGATGGCTGCTC 6796
1242 eMetGlnGlnGlnGlnGlnGlnGln..... 1252
6797 CAATGGGACAACTTGGCCAGATGGGCGAGCTGGCTAGGGGCGACAGACC 6846
1253 .....GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1265
6847 ACCCCCAATATCCAGCAGGCGCCCGCAGCAACGATCTCGACAGCAGCA 6896
1265 nGlnGlnThrGlnAlaPheSerProProProAsnValThrAlaSerPro 1282
6897 GATGAAGCAACAATTTGGGTACACAGCCAGCCGCAACCCCATGAGCCGCC 6946
1282 eMetAspGlyLeuLeuAlaGlyProThrMetProGlnAlaProProGln 1298
6947 AGCAGCACATGCTCTCAGGACAGCAGCGCTCACACATCTCCTGGCCAG 6996
1299 GlnPheProGlnProAsnTyrglyMetGly.....G1 1310
6997 CAGATGCCACATCCCTTAGTAAACAGGTGGCATCTCCAGCCCGCTGCA 7046
1310 nGlnProAspProAlaPheGlyArgValSerSerProProAsnAlaMet 1327
7047 GCTCTCAGCGCCCA.....TCCCAACCTCCACATTTCCAGCC 7084
1327 eSerSerArgMetGlyProSerGlnAsnProMetMetGlnHisProGln 1343
7085 CGTCACACAGGATACACCCCGCTTCACACACCATGTTTACCCCGAG 7134
1344 AlaAlaSerIleTyrglnSerSerGluMetLysGlyTrpProSerGly 1360
7135 ACT.....GGAACCCCTCACCTCG 7154
1360 nLeuAlaArgAsnSerSerPheSerGlnGlnPheAlaHisGlnGly 1377
7155 ACTCGAGTACCATGGCCAGCTCCCATGTGATCAG...GGACACCTGGGA 7201
1377 snProAlaValTyrsMetVal 1384
7202 ACCCTGAACAGAGTGCATGCTC 7224

seq_name: /cgn2_6/ptodata/1/lna/6A_COMB.seq:US-08-961-739-1
seq_documentation_block:
: Sequence 1, Application US/08961739A
: Patent No. 6063583
: GENERAL INFORMATION:
: APPLICANT: Montminy, Marc R.
: TITLE OF INVENTION: Methods for Treating Diabetes Mellitus
: FILE REFERENCE: SALK1650-1
: CURRENT APPLICATION NUMBER: US/08/961,739A
: EARLIER FILING DATE: 1997-10-31
: EARLIER APPLICATION NUMBER: US 194,468
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 7326
: TYPE: DNA
: ORGANISM: Mus
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)...(7326)

```

```

: NAME/KEY: misc_feature
: LOCATION: (1)...(7326)
: OTHER INFORMATION: n = A,T,C or G
US-08-961-739-1

alignment_scores:
  Quality: 254.50      Length: 591
  Ratio: 0.887         Gaps: 27
  Percent Similarity: 48.562  Percent Identity: 24.873

alignment_block:
US-09-041-994-2 x US-08-961-739-1  ..

Align seg 1/1 to: US-08-961-739-1 from: 1 to: 7344

857 SerLeuAspSerProValSerValGlySerSerProProValLysAsn11 873
||||| |||||||:|||||: |||:|||||:
5644 AGTTGCTTCTCTCTACCTAGCACCCGGGACTCTACACAGCAG.. 5691
873 eSerAlaPheProMetLeuProLysGlnProMetLeuGlyLysAsnPro 890
5692 .....CCAGCACACCCCAAAACACACAGCCCCAGCCAGCCCTC 5731
890 rGMetMetAspSerGlnGluAsnTyrglySerSerMetGlyLysProAsn 906
5732 AG.....CCTTCA 5739
907 ArgAsnValThrValThrGlnThrProSerSerGlyAspTrpGlyLeuPr 923
5740 .....CCTGTACATGTCACAGCA.....GCTTCC 5768
923 oAsnSerLysAlaGlyArgMetGluProMetAsnSerAsnSerMetGly 940
5769 TAAT.....GTAGCCGAGCTCAGCCCAACAATAGTGTGCGGGA 5812
940 rGProGlyLysPtyrAsnThrSerLeuProArgProAlaLeuGlyGly 956
5813 AGCT.....ACCAACCAAGTGCAGCTCCCGCCCGCTGCTGC 5850
957 SerIleProThrLeuProLeuArgSerAsnSerIleProGlyAlaArgPr 973
5851 CAGCCCCACCTGCA.....GCAGTAGAAGCAGCCCGGCA 5885
973 oVal.....LeuGlnGlnGlnGlnMetLeuGln..... 983
5886 AATTGAAGCTGAGGCCAGCAGCAGCAGCAGCATATCGAGCAACATCA 5935
984 .....MetArgProGlyGluIleProMetGly..... 992
5936 ACAATGGCATGCCCCCGAGCAGTACGTATGGGACCCAGAAACCA 5985
993 .....MetGlyAlaAsn.....ProTyrgly 1000
5986 ATGACTCTGTGGGCTGGAATGTGCCCGTCCCAACCAAGTCAAGTGGCC 6035
1000 nAlaAlaAlaSerAsnGlnLeuGlySerTrpProAspGlyMetLeuSer 1017
6036 TGTCAATGCTAGTATAGCCACCTGGGAGTGGAGCAGCAGCAGCAGCAG 6085
1017 eGluGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSer 1033
6086 ACAGCAGCCGATGCCAGGC...ATGCCAGGCGTATATGCCATGAC 6132
1034 LeuAspAspLeuValGlyProProSerAsnLeuGlnGlyGlnSerAsp 1050
6133 GCCCAGGACGAGTGGCTGGCCACGATCCCAATGTCCAGCAACAG 6182
1050 u.....ArgAlaLeuAspGlnLeuHisThrLeuLeuSer 1063
6183 GAGCATCTGCGCAAGTCCCTGCAGAGACCTGCTAGGAGCCCTAAAGTCA 6232
1063 snThrAspAlaThrGlyLeuGlnGluIleAspArgAlaLeu...GlyIle 1078

```

```

6233 CCAGCTCTCTCAGCAGCAGCAGCTGCTGAACATCTTAATATCAAC 6282
1079 ProGUleuValaInglInglInalaleuGUProlySlnsPalap 1095
6283 CCACAGCTATG.....GCAGCTTT 6302
1095 eGInglInglInalalalalalMetelaspInlYsAlaglyLeuTYG 1112
6303 CATCAACAGCCAGCAGCAGCAAGATGTGGCAATCAGCAGCTGCAGCAG 6352
1112 LY.....GlnThrTYProAlaInglInglProprometInglY 1124
6353 CCCAGCCCGGAGCTTCAATCCAGCTGTGATGAGCC.....CAGCCT 6396
1125 GlyPheHisLeuInglInglInserProSerPheAsnSerMetelasnGI 1141
6397 GGCATCAGCAGCAGCAGCTGATGCAAAACCTGACAGCAATGCAAGCTGG 6446
1141 nMetasnInglInglYAsnPheProleuGln.....GlyMetH 1154
6447 TGTGCCAGCGCTGTGTGCTCTCCACCACACAGCAATGGAGGCTGA 6496
1154 IsPProArGAla.....AsnIleMetArProArGThrAsn..... 1165
6497 ATCCCAAGGAGACAGCTGACATCATGACACCCAGACACACCCCAAC 6546
1166 .....ThrProlySlnLeuArGmetGlnleuInglInglY 1178
6547 ATGACCAACATGAACTCACAGTACCGAAATGTGAGAGCAGCAGCTGG 6596
1178 uGInglInglInPheleuasnInglInserArGInalaleuGlnleuLysM 1195
6597 ACAGCACCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 6646
1195 eGInlAsnProThr...AlaGlylAlalalalalMetArGprometMet 1210
6647 AAAATGTCGACAGCTGGCGGGGGGATGGCGGAGACAGCAGCTTCAG 6696
1211 GlnProGln.....GlnGlyPheleuasnAlaGlnMetValalaglnAr 1225
6697 CAGCCACAGCAGCAGCTGAGGTATGCCCCAGCATCAGCAGCAGCAGC 6746
1225 gSerArGyluleuLeuSerHisPheArGInglInglArGValalalMet 1242
6747 GCACAGCAGCTCCCATCCAGGCGAGCTCCATGAGGCCAGATGGCTGTC 6796
1242 eMetGlnInglInglInglInglInglIn..... 1252
6797 CAATGGGACAACTTGGCAGATGGGCGAGCTGAGGCGAGCAGCAGC 6846
1253 .....GlnGlnInglInglInglInglInglInglInglInglIngl 1265
6847 ACCCTAATATTCAGCAGCGCCCTGAGCAAGGATTCCTGAGCAGCAGCA 6896
1265 nGlnInglInglInAlaPheSerProProAsnValThlAlaSerProS 1282
6897 GATGAAGCAACAATTTGGCTACCCAGCAGCCAGCCGAACCCCATGAGCC 6946
1282 eMetAspGlyLeuAlaGlyProThrMetProGlnAlaProGln 1298
6947 AGCAGCAGCATGCTCTCAGAGCAGCAGCAGCAGCTCATCTCCCTGGCAG 6996
1299 GlnPheProTYGlnProAsnTYrGlyMetGly.....GI 1310
6997 CAGATCGCACATCCTTAGTAACAGGTCCAGTCTCAGCCCTGTGTA 7046
1310 nGlnProAspProAlaPheGlyArGValSerSerProProAsnAlaMet 1327
7047 GTTCCACGCGCCCA.....TCCCAACTCCACATTCAGCC 7084
1327 eSerSerArGmetGlyProSerGlnsnProMetMetGlnHisProGln 1343

```

```

7085 CGTCACACGAGATACACCCAGCTTCACACACCATGTTTCACCCAG 7134
1344 AlaAlaSerIleTYrGlnSerSerGlnMetLysGlyTYrProSerGlyAs 1360
7135 ACT.....GGAACCCCTCACCCTGG 7154
1360 nLeuAlaArGAsnSerSerPheSerInglInglInPheAlaHisInglY 1377
7155 ACTGGAGTCACCATGGCCAGCTTCATGATGATGAG...GGACACCTGGGA 7201
1377 snProAlaValTYrSerMetVal 1384
7202 ACCCTGACAGAGTGCATGCTC 7224

```

seq\_name: /cgn2\_6/ptodata/1/jna/5A\_COMB.seq:US-08-045-806-3

seq\_documentation\_block:

Sequence 3, Application US/08045806

Patent No. 5378822

GENERAL INFORMATION:

APPLICANT: Bradfield, Christopher Alan

APPLICANT: Polwick, Kristin Marie

APPLICANT: Poland, Alan

TITLE OF INVENTION: Ah Receptor cDNA and Method of

NUMBER OF SEQUENCES: 23 Determining Human Risks To Environmental Pollutants

CORRESPONDENCE ADDRESS:

ADDRESSEE: Tilton, Fallon, Lungmus & Chestnut

STREET: 100 South Wacker Drive, Suite 960

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606-4002

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/045, 806

FILING DATE: 19930408

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Rentress, Susan B.

REGISTRATION NUMBER: 31,327

REFERENCE/DOCKET NUMBER: NU-9207

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312)-456-8000

TELEFAX: (312)-456-7776

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 5261 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: double

TOPOLOGY: unknown

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 383..2927

US-08-045-806-3

alignment\_scores:

Quality: 252.00

Ratio: 0.545

Percent Similarity: 43.958

Percent Identity: 19.696

alignment\_block:

US-09-041-994-2 x US-08-045-806-3

Align seg 1/1 to: US-08-045-806-3 from: 1 to: 5261

15 SerArGlysnArGlysnProCysAsp.....ThrProGI 26



```

532 ValGlyThrLeuLeuSerThrLeuSerSerProGlyProluysuA 549
      :::::|:::
2017 CCAAGCATGATTTT..... 2032
549 spAsnSerProAsnMetAsnIleThrGlnProSerLysValSerAsnGln 565
      :::::|:::
2033 .....GAAGACATCAGACATGCAAGTGAATAATTTTCAGAAAT 2074
566 AspSerLysSerProLeuGlyPheTyrCysAspGlnAsnProValGluSe 582
      ||| :::::|:::
2075 GATTTCCTGTGTGAGTTGACTTC..... 2098
582 rSerMetCysGlnSerAsnSerArpasp...HisLeuSerAspLysGlu 598
      ||||| :::::|:::
2099 .....AGAGCATGACTTAAGCAT..... 2119
598 erLysGlnSerSerValGluGlyAlaGlnAsnGlnArgGlyProluGln 614
2119 ..... 2119
615 SerLysGlyHisLysLysLeuLeuGlnLeuThrCysSerSerAspAs 631
      :::::|:::
2120 .....GAATCTCTGACGTATGTCACAGAT.. 2143
631 pArgGlyHisSerSerLeuThrAsnSerPro..... 641
      ||||| :::::|:::
2144 .....TCATTAAAGTACTCCCTTCATACCTCAGATTATC 2180
642 .....LeuAspSerSerCysLysGlnSerSerVal 651
      ||||| :::::|:::
2181 AACAGCAACAGTCCTTGCTGTGACTCAAGCTGTATGCT..... 2220
652 SerValThrSerProSerGlyValSerSerSerThrSerGlyValSe 668
      :::::|:::
2221 ACAGAGAACACCTACATCTAGACAGCAACCAACATCACCACCAAGCAG 2270
668 rSerThrSerAsnMetHisGlySerLeuLeuGlnGluLysHisArgIle 684
      ||||| :::::|:::
2271 TAGTAGTGAGGCCACCAACCAAGCTGTGTGAGAGATGAAGACACATGCCAA 2320
685 LeuHisLysLeuLeuGlnAsnGlyAsnSerProAlaGluValAlaLysI 701
      :::::|:::
2321 GTTAATGGCATGTTGAATAATGGAATCTAACCAAAATCGTCCTTCA 2370
701 eThrAlaGlnAlaThrGlyLysAspThrSerSerIleThrSerCysGlyA 718
      ||| :::::|:::
2371 TTGTCCACAG.....CAAGACCCCAACAATATAATGTCTTTACAG 2411
718 spGlyAsnValValLysGlnGlnGlnLeuSerProLysLysGlnAsn 734
      ||| :::::|:::
2412 ACTTACCTGGGATCAGTCAAGAG.....TTCCCTTCAAAATCGAATG 2455
735 AsnAlaLeu.....LeuArgTyrLeuLeuAspArgAspAspProSerAs 749
      :::::|:::
2456 GATTCTATGCGTTATACACAGACTTATTTCCGTAAATCAGCCT.... 2500
749 pAlaLeuSerLysGlnLeuGlnProGlnValGluGlyValAspAsnLysm 766
      :::::|:::
2501 .....GTATTACCACAA.....C 2513
766 eIleGlnGlyThrSerSerThrIleProSerSerSerGlnGlnLysAsp 782
      ||||| :::::|:::
2514 ATTCCAAATATACAGAGCTGACTACCTATGGGAGTTTGAACCCATCC 2563
783 ProLysIleLysThrGlnThrSerGlnGlnGlySerGlyAspLeuAspAs 799
      ||| :::::|:::
2564 CCAATACCCCACTACTTCTAGTTTGA..... 2580
799 nLeuAspAlaIleLeuGlnLysLeuThrSerSerAspPheTyrAsn...A 815
      ||||| :::::|:::
2591 .....GATTTCCTCACTTGT 2606

```

```

815 snSerIleSerSerAsnGlySerHisLeuGlyThrLysGlnGlnValPhe 831
      :::::|:::
2607 TACACTTCCTGAAACCAAGCAT..... 2632
832 GlnGlyThrAsnSerLeuGlyLeuLysSerSerGlnSerValGlnSerI 848
      ||||| :::::|:::
2633 .....GGATTAAATCCACAGTCAAGCATTAAATCACTCC 2664
848 eArgProProTyrAsnArgAlaValSerLeuAspSerProValSerValG 865
      :::::|:::
2665 TCAGACATGTTATGCTGGGCCCGGTGATGATCACTGCCACGACGAAAC 2714
865 LysSerSerProProValLysAsnIleSerAlaPheProMetLeuProLys 881
      ||| :::::|:::
2715 CTCAGCACACCCACAGTGGGTGATGATGATCAATCCAGTACGACGCGC 2764
882 GlnProMetLeuGlyLysAsnProArgMetMetAspSerGlnGlnuSnTy 898
      ||| :::::|:::
2765 CAACAGCATTTTAAACAGATTCAAGATGAGAGTTTAAATGAACATA 2814
898 rGlySerSerMetGlyGlyProAsnArgAsnValThrValThrGlnThrP 915
      ||| :::::|:::
2815 TCCAGCTGAATTA.....ATAACATAAATATACACTCAGACTA 2852
915 rSerSerGlyAspTyrGlyLeuProAsnSerLysAlaGlyArgMetGlu 931
      :::::|:::
2853 CCACA.....CATCTCAG 2866
932 ProMetAsnSerAsnSerMetGlyArgProGlyLysPtyrAsnThrSe 948
      ||||| :::::|:::
2867 CCACTTATCATCCGTGACGAAGACGACGACCTTTTCGTGATTTCACATCCAG 2916
948 r 948
2917 T 2917

```

seq\_name: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:US-08-366-051B-3

```

seq_documentation_block:
; Sequence 3, Application US/08366051B
; Patent No. 5650283
; GENERAL INFORMATION:
; APPLICANT: Bradfield, Christopher A.
; APPLICANT: Dolwick, Kristin M.
; APPLICANT: Carver, Lucy A.
; TITLE OF INVENTION: Ah Receptor cDNAs and Genetically
; TITLE OF INVENTION: Engineered cells for Detecting Agonists to the Ah
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESS: Tilton, Fallon, Lungmus & Chestnut
; STREET: 100 South Wacker Drive, Suite 960
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-4002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/366, 051B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Tilton, Timothy L.
; REGISTRATION NUMBER: 16,926
; REFERENCE/DOCKET NUMBER: NU-9207-CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)-456-8000
; TELEFAX: (312)-456-7776
; INFORMATION FOR SEQ ID NO: 3:

```





442 .....LeuThrProGlyProGlyMetGlnSerP 451  
 1765 GCACAAAGATGAGCTGTTATCTCTATCTGCTCAAGTACTTCAAGTA 1814  
 451 roSerSerTyrgLInAsnAsnAsnTyrGlyLeuAsnMetSerSerProPro 467  
 1815 CTGCACCTTTGAAACAACTTTTTCACGAACTATGATGATGATGAGA 1864  
 468 HisGlySerProGlyLeuAlaProAsnGlnGlnAsnIleMetIleSerPr 484  
 1865 AATTGGCAAGATAATCTGCACCG..... 1888  
 484 oArgAsnArgGlySerProLysIleAlaSerHisGlnPheSerProValA 501  
 1889 .....ATGGGAATGATATCTATCTCGAAACATGAG..... 1918  
 501 IacGlyAlaHisSerProMetAlaSerSerGlyAsnThrGlyAsnHisSer 517  
 1919 ..CAAAATTGACACCTCAGATGTGAACATCATTTGCTGGAGGTCCACCA 1966  
 518 .....PheSerSerSerSerLeuSerSerAlaLeuGlnAlaIleSerGluG 532  
 1967 GGGCTCTTCAAGATAGTAAACACGACTGTGACATGATGATGAAAAA 2016  
 532 yValGlyThrSerLeuLeuSerThrIleSerSerProGlyProLysLeuA 549  
 2017 CTTAGGCAATGATTTT..... 2032  
 549 sPaSerSerProAsnMetAsnIleThrGlnProSerLysValSerAsnGln 565  
 2033 .....GAAGACATCAGACACATGCAGATGAAATTTTTCAGAAAT 2074  
 566 AspSerLysSerProLeuGlyPheTyrCysAspGlnAsnProValGluSe 582  
 2075 GATTTTCTGTGTGAGCTTGAATC..... 2098  
 582 rSerMetCysGlnSerAsnSerAspArgp...HisLeuSerAspLysGluS 598  
 2099 .....AGAGACATTTGACTTAAACGAT..... 2119  
 598 eArgLysGlnSerSerValGluGlyAlaGlnAsnGlnArgGlyProLeuGln 614  
 2119 ..... 2119  
 615 SerLysGlnHisLysLysLeuLeuGlnLeuLeuThrCysSerSerAspAs 631  
 2120 .....GAAATCCTGACGTATGTCCAAAGAT.. 2143  
 631 parGlyHisSerSerLeuThrAsnSerPro..... 641  
 2144 .....TCTTTAAGTAAGTCTCCCTTCATACCTTCAGATTATC 2180  
 642 .....LeuAspSerSerCysLysGlnSerSerVal 651  
 2181 AACGCAACAGTCTTGGCTCTGAACCTCAACCTTATGCT..... 2220  
 652 SerValThrSerProSerGlyValSerSerSerThrSerGlyGlyValSe 668  
 2221 ACAGGAACACCTACATCTAGAACAGCAACACATCACCACCAAGCAAG 2270  
 668 rSerThrSerAsnMetHisGlySerLeuLeuGlnGluLysHisArgIle 684  
 2271 TAGTAGTGGAGCCACAGCAACAGCTGTGTGAGAGATGAGACACATGCCAA 2320  
 685 LeuHisLysLeuLeuGlnAsnGlyAsnSerProAlaGluValAlaLysII 701  
 2321 GTTAATGCGCATGTTGAATAATTGAACTTAACAATAATCGCTTCA 2370  
 701 eThrAlaGlnAlaThrGlyLysAspThrSerSerIleThrSerCysGlyA 718  
 2371 TTGTCCACAG.....CAAGACCCACCAACAATAATATGCTTTACAG 2411  
 718 spGlyAsnValValLysGlnGlnLeuSerProLysLysLysGlnAsn 734

2412 ACTTACATGGATGATCACAAG.....TTCCCTCAAAATCTGAATATG 2455  
 735 AsnAlaLeu.....LeuArgTyrLeuLeuAspArgAspAspProSerAs 749  
 2456 GATTCATGCTTATATACAGAACTTATTTCTGTATATCAGCCT..... 2500  
 749 pAlaLeuSerLysGlnLeuGlnProGlnValGluGlyValAlaAspAsnLysM 766  
 2501 .....GATTACCAACA.....C 2513  
 766 eSerGlnCysThrSerSerThrIleProSerSerSerGlnGluLysAsp 782  
 2514 ATTCCAATGTACAGACCTGACATCCCTATGGGAGATTTTGAACCTCC 2563  
 783 ProLysIleLysThrGlnThrSerGlnGluGlySerGlyAspLeuAspAs 799  
 2564 CCATACCCCACTACTCTAGTTAGAA..... 2590  
 799 nLeuAspAlaIleLeuGlnLysPheLeuThrSerSerAspPheTyrAsn...A 815  
 2591 .....GATTTGTCACTTGTT 2606  
 815 snSerIleSerSerAsnGlySerHisLeuGlyThrLysGlnGlnValPhe 831  
 2607 TACAACTTCTCGAAACCAAAAGCAT..... 2632  
 832 GlnGlyThrAsnSerLeuGlyLeuLysSerSerGlnSerValGlnSerII 848  
 2633 .....GGATTAATTCACACAGTCCATCAATATCAATCCTC 2664  
 848 eArgProProTyrAsnArgAlaValSerLeuAspSerProValSerValG 865  
 2665 TCAGACATGTTATGCTGGGCGCTGATGTATCACTGCCAGCCAGCAAC 2714  
 865 LysSerSerProProValLysAsnIleSerAlaPheProMetLeuProLys 881  
 2715 CTCAGCACACCCACGCTGGGTGACATGCATCAATCCAGTACTGCCAGC 2764  
 882 GlnProMetLeuGlyGlyAsnProArgMetMetAspSerGlnLysAsnTy 898  
 2765 CAACAGCATTTTAAACAAGTTTCAGATGAGACTTTTAATGAAACATA 2814  
 898 rGlySerSerMetGlyGlyProAsnArgAsnValThrValIleGlnThrP 915  
 2815 TCCAGCTGAATTA.....AATAACATTAATATACACTCAGACTA 2852  
 915 roSerSerGlyAspTyrGlyLeuProAsnSerLysAlaGlyArgMetGlu 931  
 2853 CCACA.....CATCTTCAG 2866  
 932 ProMetAsnSerAsnSerMetGlyArgProGlyGlyAspTyrAsnThrSe 948  
 2867 CCAATTATCATATCGTCAAGAACCAAGACCTTTTCTGATTTGACATCCAG 2916  
 948 r 948  
 2917 T 2917

---



51 LeuIleSerAlaAsnLeuSerAspIleAspAsnPheAsnValLysProAs 67  
 |||||  
 236 CTGATATCTGCCAATCTTAGTATATTGCATTTCAATGTCACAAACAGA 285  
 67 pLyGcysAlaIleLeuLysGluThrValArgInIleArgInIleLysG 84  
 |||||  
 286 TAAATGTGGGATTTTAAAGGAACAGTAAGACATACGTCAAATTAAGA 335  
 84 LuGInGlySerThrIleSerAsnAspAspValGInLysAlaAspVal 100  
 |||||  
 336 AGCAAGGAAAACTATTCTCCATGATGATGATGTTCAAAAAGCCGATGA 385  
 101 SerSerThrGlyGlyGlyValIleAspLysAspSerLeuGlyProLeu 117  
 |||||  
 386 TCTTCTACAGGGCAGGAGATTATTGATTAAGACTCTTAGGACCCCTTT 435  
 117 uLeuGlnAlaLeuAspGlyPheLeuPheValValAsnArgGluAlaAsnI 134  
 |||||  
 436 ACTTACGGCATTGGATGGTTCTTATTTGGTGGATGACAGAGCAACCA 485  
 134 LeValPheValSerGluAsnValThrGlnTyrrLeuGlnTyrrLysGln 150  
 |||||  
 486 TTGTATTGTATCAGAAAAATGTCACACAATACCTGCAATATATAAGCA 535  
 151 AspLeuValAsnThrSerValTyrrAsnIleLeuHisGluLysAspArg 167  
 |||||  
 536 GACCTGGTTAACCAAGTGTTTACAAATATCTTACATGAAGAAGACAGAA 585  
 167 sAspPheLeuLysAsnLeuProLysSerThrValAsnGlyValSerTrp 184  
 |||||  
 586 GGATTTTCTTAAGAAATTACCAAAATCTACACTTAATGAGATTCTCTG 635  
 184 hrAsnGluProGlnArgGlnLysSerHisThrPheAsnCysArgMetLeu 200  
 |||||  
 636 CAAATAGAGCCCAAGACAAAAAACCCATACATTAAATGGCCGTATGTTG 685  
 201 MetLysThrProHisAspIleLeuGlnLysPheAsnAlaSerProGlu 217  
 |||||  
 686 ATGAAACACCCACATGATATTTCTGGAAGACATTAACGCCAGTCCGAAT 735  
 217 LArgGlnArgTyrrGluThrMetGlnCysPheAlaLeuSerGlnProArg 234  
 |||||  
 736 GGGCCAGAGATATGAACAATGCACTGCTTCCCTGTCTCAGCCACAGAG 785  
 234 LameTmetGluGluGluAspLeuGlnSerCysMetIleCysValAla 250  
 |||||  
 786 CTATGATGAGGAAGGGGAAGATTGCAATCTTGTATGATCTGTGGCA 835  
 251 ArgArgIleThrThrGlyGluArgThrPheProSerAsnProGlnSerPh 267  
 |||||  
 836 CGCCGCAATTACTACAGAGAAAGAACATTTCCATCAAAACCCGAGAAGCT 885  
 267 eIleThrArgHisAspLeuSerGlyLysValValAsnIleAspThrAsn 284  
 |||||  
 886 TATTCCAGACATGATCTTTTCAGAAAGTTTCTCAATATAGATACAAAT 935  
 284 erLeuArgSerSerMetArgProGlyPheGluAspIleIleArgArgCys 300  
 |||||  
 936 CACTGAGATCCCTCATGAGGCTGCTTGAAGATATATATCCGAGAGTGT 985  
 301 lIleGlnArgPhePheSerLeuAsnAspGlyGlnSerTrpSerGlnLysArg 317  
 |||||  
 986 ATTCAAGATTTTTTAAGTCTAAATGATGGCGCATGATGTCACAAACG 1035  
 317 gHisTyrrGlnGluAlaTyrrLeuAsnGlyHisAlaGluThrProValTyrrA 334  
 |||||  
 1036 TCACATACAGAAAGCTTATCTTAATGCCCATACAGAAACCCAGATATATC 1085  
 334 rgPheSerLeuAlaAspGlyThrIleValThrAlaGlnThrLysSerLys 350  
 |||||  
 1086 GATTCGTGTTGGCTGATGAACTATAGTACGACAGACAGCAAAAAGCAAA 1135

351 LeuPheArgAsnProValThrAsnAspArgHisGlyPheValSerThrHis 367  
 |||||  
 1136 CTCTCCGGAATCTGTAAACAATGATGACATGCTTGTGTCTCAACCA 1185  
 367 sPheLeuGlnArgGluGlnAsnGlyTyrrArgProAsnProAsnProValG 384  
 |||||  
 1186 CTTCCTTCAGAGAGAACAAATGATATAGACCAAAACCAATCTGTGTG 1235  
 384 lYgInGlyIleArgProPheMetAlaGlyCysAsnSerSerValGlyGly 400  
 |||||  
 1236 GACAAGGGATTAGACCACCTATGCTGGATGACACAGATTGGTAGCGGC 1285  
 401 MetSerMetSerProAsnGlnGlyLeuGlnMetProSerSerArgAlaTy 417  
 |||||  
 1286 ATGAGTATGTCCGCAACCAAGCTTACAGATGCCAGACAGAGGCCCTA 1335  
 417 rGlyLeuAlaAspProSerThrThrGlyGlnMetSerGlyAlaArgTyrrG 434  
 |||||  
 1336 TGGCTTGGCAGACCCTAGCACACAGGGCAGATGATGAGCTTAGGTATG 1385  
 434 lYgIlySerSerAsnIleAlaSerLeuThrProGlyProGlyMetGlnSer 450  
 |||||  
 1386 GGGGTTCAGTACATAGCTTATGACCCCTGGGCGCAGGCATGCATCA 1435  
 451 ProSerSerTyrrGlnAsnAsnArgTyrrGlyLeuAsnMetSerSerPro 467  
 |||||  
 1436 CCATCTCTTACAGAACAAACAACTATGGGCTCAACATGATGATGAGCC 1485  
 467 oHisGlySerProGlyLeuAlaProAsnGlnGlnAsnIleMetLieser 484  
 |||||  
 1486 ACATGGAGTCTGTGTTGCCCAACACAGCAGAAATCATGATTTCTC 1535  
 484 roArgAsnArgGlySerProLysIleAlaSerHisGlnPheSerProVal 500  
 |||||  
 1536 CTGTTAATCGTGGAGATCCAAAGATAGCCTCACATGACTTTCTCTGTT 1585  
 501 AlaGlyValHisSerProMetAlaSerSerGlyAsnThrGlyAsnHisSe 517  
 |||||  
 1586 GCAGGTGTGCACTCTCCATGGCATCTTCTGGCAATACTGGGAACCAAG 1635  
 517 rPheSerSerSerSerLeuSerAlaLeuGlnAlaIleSerGluGlyValG 534  
 |||||  
 1636 CTTTTCACAGACCTCTCTCAGTGGCCCTGCMAGCCATCTCAAGTAGTGG 1685  
 534 lYThrSerLeuLeuSerThrLeuSerSerProGlyProLysLeuAspAsn 550  
 |||||  
 1686 GGACTTCCCTTTATCTACTGTGTCATCACAGGCCCAAAATTTGATTAAC 1735  
 551 SerProAsnMetAsnIleThrGlnProSerLysValSerAsnGlnAspSe 567  
 |||||  
 1736 TCTCCCAATATGAATATTAACCCAAAGTAAGTAAGCAATCAGAGATTG 1785  
 567 rLysSerProLeuGlyPheTyrrCysAspGlnAsnProValGlnSerSer 584  
 |||||  
 1786 CAAGAGTCTCTGGGCTTTTATTTGCGACCAAAATCAAGTGAAGATTCAA 1835  
 584 eCyGlnSerAsnSerArgAspHisLeuSerAspLysGluSerLysGlu 600  
 |||||  
 1836 TGTGTACAGTCAAAATAGCAGAGATCACCTCAGTACAGAAAGTAAGAG 1885  
 601 SerSerValGluGlyAlaGluAsnGlnArgGlyProLeuGlnSerLysGln 617  
 |||||  
 1886 AGCAAGTGTAGAGGGGAGAGAAATCAAAAGGGGTCTTTGGAAAGCAAAAG 1935  
 617 yHisLysLysLeuLeuGlnLeuLeuThrCysSerSerAspAspArgGlyH 634  
 |||||  
 1936 TCATATAAAATTAAGTCAAGTTACTTACTGTCTTCTGTATGACCGGGTTC 1985  
 634 lAsSerSerLeuThrAsnSerProLeuAspSerSerCysAlaGlnSerSer 650  
 |||||  
 1986 ATTCTCTCTTGACCAACTCCGCCCTAGATTCAAAGTTGAAGAAATCTTCT 2035  
 651 ValSerValThrSerProSerGlyValSerSerSerThrSerGlyGlyVal 667

```

|||||
2036 GTTAGTGTACACAGCCCTCTGGAGTCTCTCTCTACATCTGGAGAGCT 2085
|||||
667 lSerSerThrSeramethisglYSerLeuLeuInglulysHlsargI 684
|||||
2086 ATCTGTACATCCAAATATGATGGGTGCTACTGTTCAGAGAGAACGCCGA 2135
|||||
684 lEleuHlsYsLeuLeuIngluInsnclYasnSerProAlaGluValAlaLys 700
|||||
2136 TTTTGGCAAAAGTTGCTCGAATGGAATTCACACGCTGAGGTAGCCAG 2185
|||||
701 lIleThrAlaGlnAlaThrnglyLysAspThrSerSerIleThrSerCysgl 717
|||||
2186 ATTACTGTACAGACCCACTGGGAAAGACACAGCATATACCTTCTGTGG 2235
|||||
717 yAspGlyAsnValAlaLysglIngluInleuSerProLysLysglua 734
|||||
2236 GGACGGAAATGTTGTCAGAGGAGAGAGCTAAGTCTAAGAAAGAGGA 2285
|||||
734 snAsnAlaLeuLeuArgTyrLeuLeuAspArgAspProSerAspAla 750
|||||
2286 ATAATGCACTTCTTAGATACCTGCTGACAGGAGATGATCCTAGTAGATGA 2335
|||||
751 leuSerLysgluLeuGlnProGlnValAlaGluAlaAspAsnLysMetSe 767
|||||
2336 CTCTCTAAAGACTACAGCCCAAGGTGGAAGAGTGCACAAATAAATGAG 2385
|||||
767 rGlnCysThrSerSerThrIleProSerSerSerGlnlulYasnAspProL 784
|||||
2386 TCAGTGCACAGCTCCACCATTTCTAGCTCAAGTCAAGAGAAAGACCCTA 2435
|||||
784 yslleLysThrGlnThrSerGlnGluLysSerGlyAspLeuAspAsnleu 800
|||||
2436 AAATTAAGACAGAGACAAGTGAAGAGGATCTGAGACTTGGAATATCTTA 2485
|||||
801 AspAlaIleLeuGlyAspLeuThrSerSerAspPheTyrAsnAsnSerI 817
|||||
2486 GATGCTATTTCTTGCTGATCTGACTAGTTCTCTACATTATACATATTCAT 2535
|||||
817 eSerSerAsnGlySerHlsLeuGlyThrLysGlnGlnValPheGlnGlyT 834
|||||
2536 ATTCCTCAATAGTATCATCTGGGAGCTAAGCAACAGGTTTCAAGAA 2585
|||||
834 hrAsnSerLeuGlyLeuLysSerSerGlnSerValGlnSerIleArgPro 850
|||||
2586 CTAAATTTCTGGGTTTGAAGGTTGACAGTCTGACAGTCTATTCGTCCT 2635
|||||
851 ProTyrAsnArgAlaValSerLeuAspSerProValSerValGlySerSe 867
|||||
2636 CCAATATAACCGAGCAGTCTCTGTGATAGCCCTGTTTCTGTGGCTCAG 2685
|||||
867 rProProValLysAsnIleSerAlaPheProMetLeuProLysGlnProm 884
|||||
2686 TCCTCCAGTAAATAATATACAGTCTTCCCATGTATCAACAAGCAACCA 2735
|||||
884 eLleuGlyGlyAsnProArgMetMetAspSerGlnGluAsnTyrGlySer 900
|||||
2736 TGTGGGGTGGGAATCCAGAAATGATGATAGTACGGAATAATTAAGCTCA 2785
|||||
901 SerMetGlyGlyProAsnArgAsnValThrValThrGlnThrProSerSe 917
|||||
2786 AGTATGGGTGGCCAAACCGAAATGTGACTGTGACTCAGACTCCTTCTCTC 2835
|||||
917 rGlyAspTpgIyleuProAsnSerLysAlaGlyArgMetGluProMetA 934
|||||
2836 AGGAGACTGGGGCTTACCAAACTCAAAAGCGGAGAGATGAGAACCTATGA 2885
|||||
934 snSerAsnSerMetGlyArgProGlyArgTyrAsnThrSerLeuPro 950
|||||
2886 ATTCAAACTCCATGGGAAGACCAAGAGAGATTAATATCTTTTATCC 2935
|||||
951 ArgProAlaLeuGlyGlySerIleProThrLeuProLeuArgSerAsnSe 967
|||||
2936 AGACCTGCACTGGGTGGCTATATCCACATTTGCCCTCTGGGCTAATAG 2985
|||||
967 rIleProGlyAlaArgProValLeuGlnGlnGlnGlnMetLeuGlnm 984
|||||
2986 CATACAGAGTGCAGACAGCAGATTTGCAACACAGACAGACAGATGCTTCAAA 3035
|||||
984 eTargProGlyGluIleProMetGlyMetGlyAlaAsnProTyrGlyGln 1000
|||||
3036 TGAGGCCCTGGTGAATATCCCATGGGAATGGGGGCTTAATCCCTATGGCCAA 3085
|||||
1001 AlaAlaIleSerAsnGlnLeuGlySerTyrProAspGlyMetLeuSerMe 1017
|||||
3086 GCACAGCATTTAAACCAACTGGGTTCTGCGCCCATGAGCATGTTGTCAT 3135
|||||
1017 tGluGlnValSerHlsGlyThrGlnAsnArgProLeuLeuArgAsnSerI 1034
|||||
3136 GGAACAAGTTTCTCATGAGCACTCAAAATAGAGCCTTCTTAGAATGCC 3185
|||||
1034 euAspAspLeuValGlyProProSerAsnLeuGlnGlyGlnInserAspGlu 1050
|||||
3186 TGATATGATCTTGTGGGCCACCTTCCAACTGGAAGGCCAGAGTGAAGAA 3235
|||||
1051 ArgAlaLeuLeuAspGlnLeuHlsThrLeuLeuSerAsnThrAspAlaTh 1067
|||||
3236 AGACCATTTATGGACCACTGACACATCTTCTCAGCAACACAGATGCGAC 3285
|||||
1067 rGlyLeuGlnGluIleAspArgAlaLeuGlyIleProGluLeuValAsnG 1084
|||||
3286 AGGCCTGGAAAGAAATTGACAGAGCTTGGGCATCTGMACTGTCAATC 3335
|||||
1084 lnglyGlnAlaLeuGlnProLysGlnAspAlaPheGlnGlyGlnAla 1100
|||||
3336 AGGACAGGCAATTAGAGCCAAACAGAGATGCTTTCAGAGGCCAAGAAACA 3385
|||||
1101 AlaValMetMetAspGlnLysAlaGlyLeuTyrGlyGlnThrTyrProAl 1117
|||||
3386 GCATATATGATGATGATCAGAAAGCAGGATTATATGACAGACATACCCAGC 3435
|||||
1117 gGlnGlyProProMetGlnGlyGlyPheHlsLeuGlnGlyGlnSerProS 1134
|||||
3436 ACAGGGGCTTCAATGCAAGAGAGCTTTCATCTTCAGAGCAACATCACCAT 3485
|||||
1134 erPheAsnSerMetMetAsnGlnMetAsnGlnGlnGlyAsnPhProleu 1150
|||||
3486 CTTTAACTCTATGATGATCAGATGAACACAGCAAGGCAATTTTCTCTCTC 3535
|||||
1151 GlnGlyMetHisProArgAlaAsnIleMetArgProArgThrAsnThrPr 1167
|||||
3536 CAAGGAATGCACCCAGCAGCAACATCATGAGACCCCGCAAAACACCCC 3585
|||||
1167 oLysGlnLeuArgMetGlnLeuGlnGlnArgLeuGlnGlyGlnPheL 1184
|||||
3586 CAAGCAACTTGAATGACAGCTTTCAGCAGAGGCTCAGAGGCCAGAGTTT 3635
|||||
1184 euAsnGlnSerArgGlnAlaLeuGlnLeuLysMetGluAsnProThrAla 1200
|||||
3636 TGAATCAGAGCCGAGCAGGCACTTGATGAATGAAATGGAATCCCTACGTCT 3685
|||||
1201 GlyGlyAlaAlaValMetArgProMetMetGlnProGlnGlnGlyPheL 1217
|||||
3686 GGTGTGCTGGGTGATGAGGCTATATGATGACGCCCAAGAGGGTTTCTT 3735
|||||
1217 uAsnAlaGlnMetValAlaGlnArgSerArgGluLeuLeuSerHlsHsP 1234
|||||
3736 TTAATGCTCAAAATGCTGCGCCCAACGACAGAGACTGCTAAGTCAATCCT 3785
|||||
1234 hearGlnGlnArgValAlaMetMetMetGlnGlnGlnGlnGlnGln 1250
|||||
3786 TCCGCAACAGAGGGGTGCTATGATGATGACGACGACCAACAGACAGAG 3835
|||||
1251 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1267
|||||
3836 CAGCAGCAGCAGCAGCAGCAACAGCAACAGCAACAGCAGCAACAGCA 3885

```

```
1267 nThrglnAlaPheSerProProAsnValThrAlaSerProSerMetA 1284
|||||
3886 AACCCAGGCTTCAGCCACCTCTAATGTGACTGCTCCACGATGG 3935
1284 sRgLYleuLeuAlaGlyProThrMetProGlnAlaProProGlnGlnPhe 1300
|||||
3936 ATGGGCTTTGGAGGACCCACAAATGCCAAAGCTCTCCGCAACAGTTT 3985
1301 ProTYrGlnProAsnTYrGlyMetGlyGlnGlnProAspProAlaPheG1 1317
|||||
3986 CCATATCAACCAATATATGGAAATGGGCAACAACGATCCAGCCTTTGG 4035
1317 yArGylSerSerProProAsnAlaMetSerSerArGmetGlyProS 1334
|||||
4036 TCGAGTGTCAGTCCCTCCCAATGCAATGATGTCTCAAGAAATGGGTCCT 4085
1334 eRgAlaAsnProMetMetGlnHisProGlnAlaAlaSerLetyrGlnSer 1350
|||||
4086 CCCAGAAATCCATGATGCAACACCCGACGCTCATCTATCAATGATGC 4135
1351 SerGluMetLysGlyTYrProSerGlyAsnLeuAlaArgAsnSerSerPh 1367
|||||
4136 TCAGAAATTAAGGCTGGCCATCAGAAATTTGGCCAGGACAGCTCTT 4185
1367 eSerGlnGlnGlnPheAlaHisGlnGlyAsnProAlaValTYrSerMetV 1384
|||||
4186 TTCACAGACGAGATTGCCACACAGGGAATCTCGCAGTGATATGATGG 4235
1384 aLHisMetAsnGlySerSerGlyHisMetGlyGlnMetAsnMetAsnPro 1400
|||||
4236 TCACACTGTGATGGACAGAGTGTACATGATGGACAGATGAACATGAACCCC 4285
1401 MetProMetSerGlyMetProMetGlyProAspGlnLysTYrCys 1415
|||||
4286 ATGCCATGTCGGCATGGCTATGGTCTGTATCAAAATACTGCG 4330

seq_name: /cgn2_6/p/todata/1/pna/US095A_COMB.seq:US-09-513-066-1

seq_documentation_block:
: Sequence 1, Application US/09513066
: GENERAL INFORMATION:
: APPLICANT: Chen, J. Don
: APPLICANT: Leo, Christopher
: APPLICANT: Li, Hui
: TITLE OF INVENTION: NOVEL USES FOR THE RAC3 TRANSCRIPTIONAL COACTIVATOR OF
: FILE REFERENCE: UMG-026CP
: CURRENT APPLICATION NUMBER: US/09/513,066
: PRIOR APPLICATION NUMBER: USSN 09/041,994
: PRIOR FILING DATE: 1998-03-13
: PRIOR APPLICATION NUMBER: USSN 60/073,674
: PRIOR FILING DATE: 1998-02-04
: NUMBER OF SEQ ID NOS: 37
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 4496
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (86)..(4330)
US-09-513-066-1

alignment_scores:
Quality: 7375.00 Length: 1415
Ratio: 5.216 Gaps: 0
Percent Similarity: 99.929 Percent Identity: 99.929

alignment_block:
US-09-041-994-2 x US-09-513-066-1 ..
```

```
Align seg 1/1 to: US-09-513-066-1 from: 1 to: 4496
1 MetSerGlyLeuGlyGlnAsnLeuAspProLeuAlaSerAspSerArgLy 17
|||||
86 ATGAGTGGAATTAGGAGAAACCTTGATCCATCGCCAGATTCACAGAAA 135
17 sArgLYleuProCYsAspThrProGlyGlnGlyLeuThrCYsSerGlyG 34
|||||
136 ACGCAAAATGGCATGTGATATCTCCAGGACAAAGTCTTACCTCAGAGTGG 185
34 LuLsArGArGArGArGluGlnGlnSerLysTYrLLeGluLuLeuAlaGlu 50
|||||
186 AAAAAGGAGAGCGGAGCAGAAAGTAATTAATTCAAAGAAATTTGGCTGAG 235
51 LeuLLeSerAlaAsnLeuSerAspLLeAsPAsnPhAsnValLysProAs 67
|||||
236 CTGATATCTGCCAATCTTATGTATATTCACAAATTTCAATGTCAAAACGCA 285
67 pLYsCYsAlaLeuLeuLYsGluThrValArgGlnLeaArgGlnLeLYsG 84
|||||
286 TAAATGTGGGATTTTAAAGCAAAAGTAAAGACAGATACGTCAAAATPAAAG 335
84 LuGlnGlyLYsThrLLeSerAsnAspAspValGlnLYsAlaAspVal 100
|||||
336 ACGCAAGAAAAACTATTTCCAATGATGATGATGTTCAAAAACCCGATGTA 385
101 SerSerThrGlyGlnGlyValLLeAspLYsAspSerLeuGlyProLeu 117
|||||
386 TCTTTCTACGGCGGAGGACTTATTCATTAAGACTCTTAGACCCGCTTTT 435
117 uLeuGlnAlaLeuAspGlyPheLeuPheValValAsnArgGlnAlaAsnI 134
|||||
436 ACTTCAGGCAATTTGAATGGTTCTTATTTGGTGAATCAGAGGCAAAACA 485
134 LeValPheValSerGluAsnValThrGlnTYrLeuGlnTYrLYsGlnGlu 150
|||||
486 TTGTATTGTATTCAGAAATGTACACCAATACCTGCAAAATTAAGCAAGG 535
151 AspLeuValAsnThrSerValTYrAsnLLeuHisGlnGluAspArgLY 167
|||||
536 GACCTGGTTAAACCAAGTCTTACAAATATCTTACATGAAGAAAGACAGAAA 585
167 sAspPheLeuLYsAsnLeuProLYsSerThrValAsnGlyValSerTrpT 184
|||||
586 GGATTTTCTTAAGAAATTTACCAAAATCTACAGTTATGAGATTCTCTGGA 635
184 hTrAsnGluProGlnArgGlnLYsSerHisThrPheAsnCYsArgMetLeu 200
|||||
636 CAAATGAGACCCAAAGACAAAAAGCCATACATTTAATTTGCCGTATGTTGG 685
201 MetLYsThrProHisAspLLeuGlnAspLLeAsnAlaSerProGluMe 217
|||||
686 ATGAAGAAACACCAATGATATTTCTGGAAGACATRAAACGCCAGCTCGAAT 735
217 LaRgGlnArGTYrGluThrMetGlnCYsPheAlaLeuSerGlnProArgA 234
|||||
736 GCGCCAGAGATATGAACAATCAGAGCTTCCCTGCTCAGCACAGAG 785
234 lAsMetMetGlnGluGlyLysAspLeuGlnSerCYsMetLLeCYsValAla 250
|||||
786 CTATGATGGAGGAGGGAAGATTTCAAATCTTGATGATCTGTGTGGCA 835
251 ArgArGlyLeThrThrGlyLysArgThrPheProSerAsnProGluSerPh 267
|||||
836 CCGCCGATTAACACAGAGAAAGAACATTTCCATCAAAACCTGAGAGCTT 885
267 eLLeThrArGHisAspLeuSerGlyLYsValValAsnLLeAspThrAsnS 284
|||||
886 TATTTACACAGACATGATCTTTACGAAAGGTTGTCAATATATATCAAAAT 935
284 eRLeuArGSerSerMetArGProGlyPheGluAspLLeLeaRgArgCYs 300
|||||
```

936 CACTGAGATCCTCATGAGGCTGGCTTTCAGATATATATCCGAGGCT 985  
 301 ILeGlnArgPhePheSerLeuAsnAspGlyInSerTrpSerGlnLys 317  
 986 ATTCACAGATATTTTATAGTAAATGATGGGCACTAGTGTCCAGAAACG 1035  
 317 gHsTrpGlnGluAlaTrpLeuAsnGlnHisAlaGlnThrProValTyr 334  
 1036 TCACATCAAGAAAGCTTATCTTAATGGCCATGCAGAAACCCAGTATATC 1085  
 334 rPheSerLeuAlaAspGlyThrIleValThrAlaGlnThrLysSerLys 350  
 1086 GATTCTCGTGGCTGATGGAATAGTAGCTGACAGACAGACAAAGCAA 1135  
 351 LeuPheArgAsnProValThrAsnAspArgHisGlyPheValSerThrH 367  
 1136 CTCTCCGAATCCTGTAAACAAATGATCGACATGGCTTTGTCAACCCA 1185  
 367 sPheLeuGlnArgGluGlnAsnGlyTyrArgProAsnProAsnProValG 384  
 1186 CTTCCTTCAGAGAAACAGATGATATAGACCAACCAATCTCTGTTG 1235  
 384 LysGlnGlyIleArgProProMetAlaGlyCysAsnSerSerValGlyGly 400  
 1236 GACAAGGAGATTAGACCACCTATGCTGATGATGCAACAGTTGCTAGGCGGC 1285  
 401 MetSerMetSerProAsnGlnGlyLeuGlnMetProSerSerArgAlaTyr 417  
 1286 ATGAGTATGTCCGCAACCAAGGCTTACAGATCCCGAGCAGAGGCGCTA 1335  
 417 rGlyLeuAlaAspProSerThrThrGlyGlnMetSerGlyAlaArgTyrG 434  
 1336 TGGCTTGCGCAGACCTAGACACAGGCGAGATGAGTGGAGCTAGATG 1385  
 434 LysLysSerSerAsnIleAlaSerLeuThrProGlyMetGlnSer 450  
 1386 GGGGTTCCAGATAACATAGCTTACCTTGACCCCTGGCGCAGCATGCAATCA 1435  
 451 ProSerSerTyrGlnAsnAsnTyrGlyLeuAsnMetSerSerProTrp 467  
 1436 CCATCTTCCACAGAACCAACATAGGCTCAACATGATGAGTCCGCC 1485  
 467 oHsGlySerProGlyLeuAlaProAsnGlnGlnAsnIleMetIleSerP 484  
 1486 ACATGGGAGTCCGTGCTTGTCCCAACACAGCAGAAATCATGATTTCTC 1535  
 484 rOArgAsnArgGlySerProLysIleAlaSerHisGlnPheSerProVal 500  
 1536 CTCGTATCTGTGGAGTCCAAAGATAGCTCAGATCAGTTTCTCCCTGTT 1585  
 501 AlaGlyValHisSerProMetAlaSerSerGlyAsnThrGlnLysnHisSe 517  
 1586 GCAGGtGTGCATCTCCATGCGATCTTGTGCATAATCTGGAAACCCAG 1635  
 517 rPheSerSerSerSerLeuSerAlaLeuGlnAlaIleSerGlyGlyValG 534  
 1636 CTTTTCAGAGAGCTCTCTCAGTGCCTCGCAAGCATCAGTAAGAGTGTGG 1685  
 534 LysThrSerLeuLeuSerThrLeuSerSerProGlyProLysLeuAspAsn 550  
 1686 GGACTTCCCTTTTATCTACTCTCATCACAGGCCCCCAAAATTGGATTAAC 1735  
 551 SerProAsnMetAsnIleThrGlnProSerLysValSerAsnGlnAspSe 567  
 1736 TCTCCCAATATGAAATATACCAACCAAGTAAGTAAGCAATCAGGATTC 1785  
 567 LysSerProLeuGlyPheTyrCysAspGlnAsnProValGlyUserSerM 584  
 1786 CAAGAGTCCCTCTGGCTTTTATTGCGACCAAAATCCAGTGGAGAGTTCAA 1835  
 584 eTcysGlnSerAsnSerArgAspHisLeuSerAspLysGlnSerLysGlu 600  
 1836 TGTGTGCTCAAAATAGCAGAGATCACCTCAGTGCACAAAGAAATGAGGAG 1885

601 SerSerValGlnGlyAlaGlnAsnGlnArgGlyProLeuGlnSerLysG 617  
 1886 AGCAGTGTGAGGGGCGAGAAATCAAAAGGGTCTTTGGAAAGCAAAAGG 1935  
 617 yHsLysLysLeuLeuGlnLeuLeuThrCysSerSerSerAspArgGlyH 634  
 1936 TCATAAAAAATTACTGCAGTACTTACTCTCTCTCGATGAGACCGGGGTC 1985  
 634 LysSerSerLeuThrAsnSerProLeuAspSerSerCysLysGlnSerSer 650  
 1986 ATTCCTCTCTGAGCAACTCCCTCCAGATTCAGATTGTAAAGATCTTCT 2035  
 651 ValSerValThrSerProSerGlyValSerSerSerThrSerLysGlyVal 667  
 2036 GTTAGTGTACACAGCCCTCTGGAGTCTCTCTCTACATCTGGAGAGT 2085  
 667 LysSerThrSerAsnMetHisGlySerLeuLeuGlnGlnLysHisArgI 684  
 2086 ATCTCTACATCCAAATATGATGATGGTCTTACCAAGAGAACACCCGGA 2135  
 684 ILeuHisLysLeuLeuGlnAsnGlyAsnSerProAlaGlnValAlaLys 700  
 2136 TTTTGCACAAGTGTCTGCAGAAATGGCAATTCACACGCTGAGTACGCAAG 2185  
 701 IleThrAlaGlnAlaThrGlyLysAspThrSerSerIleThrSerCysG 717  
 2186 ATTAATGACACAAAGCCACTGGGAAAGACACACAGATTAATTAATCTTGTGG 2235  
 717 YAspGlyAsnValValLysGlnGlnLeuLeuSerProLysLysGlnVal 734  
 2236 GAGAGGAAATGTTGTCAAGCAGAGAGCGATAGTCTTAATAAAGAGAGA 2285  
 734 sAsnAlaLeuLeuArgTyrLeuLeuAspArgAspProSerAspAla 750  
 2286 ATATGACACTTCTTACATCTGCTGACAGGATGATCTCTAGTATGATCA 2335  
 751 LeuSerLysGlnLeuGlnProGlnValGlnGlyValAspAsnLysMetSe 767  
 2336 CTCTCTAAAGAACTACAGCCCAAGTGGAAAGAGTGGACATTAATAATGAG 2385  
 767 rGlnCysThrSerSerThrIleProSerSerSerGlnGlnLysAspProL 784  
 2386 TCATGTGACAGAGTCCACCATCTCTAGCTCAAGTCAAGAAAGACCCCTA 2435  
 784 YsIleLysThrGlnThrSerGlnGlnGlySerGlnLysLeuAspAsnLeu 800  
 2436 AAATTAAGACAGAGACAAAGTGAAGAGGATCTGAGACTTGGATTAATCTA 2485  
 801 AspAlaIleLeuGlnLysLeuThrSerSerAspPheTyrAsnAsnSerI 817  
 2486 GATGCTATTTCTTGTGTGATCTGACTAGTCTGACTTTTACATTAATTCAT 2535  
 817 eSerSerAsnGlySerHisLeuGlyThrLysGlnGlnValPheGlnGlyT 834  
 2536 ATCTCTAAATGATATGATCTGAGGAGATAAGCAACAGTCTTTCAGGAA 2585  
 834 hAsnSerLeuGlyLeuLysSerSerGlnSerValGlnSerIleArgPro 850  
 2586 CTAAATCTCTGGGTTTGAAAAGTTTCAACAGTCTGACGTATATTCGCCCT 2635  
 851 ProTyrAsnArgAlaValSerLeuAspSerProValSerValGlySerSe 867  
 2636 CCATATTAACCGAGCAGTGTCTGTGATAGCCCTGTTTCTGTGGCTCAAG 2685  
 867 rProProValLysAsnIleSerAlaPheProMetLeuProLysGlnProm 884  
 2686 TCTCTCCAGTAAAAAATATCAAGTGTCTTCCCATGTTTACCAAGACCAACCA 2735  
 884 eLeuGlnGlyLysAsnProArgMetMetAspSerGlnGlnLysTyrGlySer 900  
 2736 TGTGGGTGGGAATCCAAAGAAATGATGATGATCAGGAAATTTATGGCTCA 2785

```

901 SerMetGlyGlyProAsnArgAsnValThrValThrGlnThrProSerSe 917
|||||
2786 AGTATGGGTGGGCCCAACCGAATGTGACTGTGACTCAGACTCCTTCTCT 2835
917 rGlyAspTrpGlyLeuProAsnSerLysAlaGlyArgMetGluProMetA 934
|||||
2836 AGGAGACTGGGGCTTACCAACTCAAGCCGGCAGAAATGGAACCTATGA 2885
934 snSerAsnSerMetGlyArgProGlyLysAspTyrAsnThrSerLeuPro 950
|||||
2886 ATTCAAACTCCATGGGAGAACCGAGAGATTAATTAATCTTCTTACC 2935
951 ArgProAlaLeuGlyGlySerLleProThrLeuProLeuArgSerAsnSe 967
|||||
2936 AGACCTGCACAGGGTGGCTTATTCACATTCGCTTCCGTTCAATAG 2985
967 rLleProGlyAlaArgProValLeuGlnGlnGlnGlnMetLeuGlnM 984
|||||
2986 CATACCAGGTGGAGACAGTATTCACACAGCAGCAGAGATGCTTCAA 3035
984 eArgProGlyGlyLleProMetGlyMetGlyAlaAsnProTyrGlyGln 1000
|||||
3036 TGAGGCTGTGTAATAATCCCATGGGAATGGGGCTAATCCCTATGGCCA 3085
1001 AlaAlaAlaSerAsnGlnLeuGlySerTrpProAspGlyMetLeuSerMe 1017
|||||
3086 GCAGCAGCATCTTAACCAACTGGGTTCCTGGCCCGATGGCATGTGTCCAT 3135
1017 tGluGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerL 1034
|||||
3136 GGAACAATTTCTCATGGCAGCTCAAAATAGGCCCTTCTTAAGAAATTC 3185
1034 euAspAspLeuValGlyProProSerAsnLeuGlyGlnSerAspGlu 1050
|||||
3186 TGATGATCTTTTGGGCCACCTTCCAACTGGAAGGCCAGAGTACGAA 3235
1051 ArgAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaTh 1067
|||||
3236 AGAGCATTTATGGACACAGCTGCACACTCTTCTCACCAACACAGATGCC 3285
1067 rGlyLeuGluGluLleAspArgAlaLeuGlyLleProGluLeuValAsnG 1084
|||||
3286 AGGCTGGGAAGAAATGACAGAGCTTTGGGCACTTCCTGAACTTGTCAATC 3335
1084 lngGlyAlaLeuGluLupProLysGlnAspAlaPheGlnGlyGlnGluA 1100
|||||
3336 AGGACAGGCAATTAGACCCAAACAGATGCTTCCAAAGGCCAAGAGCA 3385
1101 AlaValMetMetAspGlnLysAlaGlyLeuTyrGlyGlnThrTyrProAl 1117
|||||
3386 GCAGTAATGATGGATCAGAAAGCAGATTAATATGACAGACATACCACAC 3435
1117 agGlnGlyProProMetGlnGlyGlyPheHisLeuGlnGlnGlnSerProS 1134
|||||
3436 ACAGGGGCTCCAAATGCAGAGAGGCTTTCAATCTTCAAGGACACATACCAT 3485
1134 ePheAsnSerMetMetAsnGlnMetAsnGlnGlnGlnGlnGlnGlnGln 1150
|||||
3486 CTTTAACTCTTATGATGATCATGATGAACACAGCAAGCAATTTCTCTC 3535
1151 GlnGlyMetHisPheProArgAlaAsnLleMetArgProArgTrpAsnThrP 1167
|||||
3536 CAAGGAATGACCCACAGGCCAACAATCATGAGAACCCGACAAACACCCC 3585
1167 oLysGlnLeuArgMetGlnLeuGlnGlnArgLeuGlnGlnGlnGlnPheL 1184
|||||
3586 CAAGCAACTTAAATATGCAGCTTCAAGCAGAGGCTGCAAGGCCACGAGTTT 3635
1184 euAsnGlnSerArgGlnAlaLeuGlnLeuLysMetGluAsnProThrAla 1200
|||||
3636 TGAATCAGAGCGCAGACAGGCACTTGAATTGAAATGAAACCTTACTGCT 3685
1201 GlyGlyAlaAlaValMetArgProMetMetGlnProGlnGlnGlnPheLe 1217

```

```

|||||
3686 GGTGTGCTGCGGTATGATGAGGCTTATGATGACAGCCAGAGGGTTTCT 3735
1217 uAsnAlaGlnMetValAlaGlnArgSerArgGluLeuLeuSerHisLSP 1234
|||||
3736 TAATGCTCAAAAGGTGCGCCAAACGACAGAGAGAGTGTGAATGTCACCT 3785
1234 hearGlnGlnArgValAlaMetMetMetGlnGlnGlnGlnGlnGlnGln 1250
|||||
3786 TCCGACAAACAGAGGGTGGCTATGATGATGACAGCAGACACACAGCAG 3835
1251 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1267
|||||
3836 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3885
1267 nThrGlnAlaPheSerProProProAsnValThrAlaSerProSerMetA 1284
|||||
3886 AACCCAGGCTTTCAGCCCACTCTTAATGTGACTCTTCCCCAGCAGTGG 3935
1284 sPGLyLeuLeuAlaGlyProThrMetProGlnAlaProProGlnGlnPhe 1300
|||||
3936 ATGGGCTTTTGGCAGACCCCAACATGCCAAGTCTCTCCGACACATTT 3985
1301 ProTyrGlnProAsnTyrGlyMetGlyGlnGlnProAspProAlaPheG 1317
|||||
3986 CCATATCAACCAAAATTAATGAAATGGAGACACACACAGATCCAGCTTGG 4035
1317 yArgValSerSerProProAsnAlaMetMetSerSerArgMetGlyProS 1334
|||||
4036 TCGAGTGTCTAGTCTCCCAATGCAATGATGTGTGCAAGAAATGGGTCTCT 4085
1334 eArgLAsnProMetMetGlnHisProGlnAlaAlaSerLleTyrGlnSer 1350
|||||
4086 CCCAGAAATCCCATGATGCAACACCCGACAGGCTGCATTCATCATACATC 4135
1351 SerGluMetLysGlyTrpProSerGlyAsnLeuAlaArgAsnSerSerPh 1367
|||||
4136 TCAGAAATGAAAGGCTGGCCATCAGAGAAATTTGGCCAGAGAACGCTCTT 4185
1367 eSerGlnGlnGlnPheAlaHisGlnGlyAsnProAlaValTyrSerMetV 1384
|||||
4186 TTCCACAGCAGCACTTGGCCACACAGGCAATCTCAGATGATATATG 4235
1384 aHisMetAsnGlySerSerGlyHisMetGlyGlnMetAsnMetAsnPro 1400
|||||
4236 TGCACATGAAATGACAGAGTGTGCATATGGACATGAAATGAACCC 4285
1401 MetProMetSerGlyMetProMetGlyProAspGlnLysTyrCys 1415
|||||
4286 ATGCCATGTCTGGCATGCTTAATGAGTCTCTGATCAGAAATTAATCTGC 4330
seq_name: /cgn2_6/ptodata/1/pna/us091_COMB.seq.us-09-125-635-1
seq_documentation_block:
; Sequence 1, Application US/09125635
; GENERAL INFORMATION:
; APPLICANT: THE UNITED STATES OF AMERICA represented by THE SE
; TITLE OF INVENTION: A1b1, A novel steriod receptor co-activator
; FILE REFERENCE: 49944
; CURRENT APPLICATION NUMBER: US/09/125,635
; PRIOR FILING DATE: 1998-08-21
; PRIOR APPLICATION NUMBER: 60/049,728
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 6835
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(4463)
US-09-125-635-1

```



## alignment\_scores:

Quality: 7352.50 Length: 1420  
 Ratio: 5.200 Gaps: 1  
 Percent Similarity: 99.577 Percent Identity: 99.366

## alignment\_block:

US-09-041-994-2 x US-09-125-635-1 ..

Align seg 1/1 to: US-09-125-635-1 from: 1 to: 6835

```

1 MetSerGlyLeuGlyGluAsnLeuAspProLeuAlaSerAspSerArgly 17
  |||||
201 ATGAGTGGATTAGAGAAACTTGATGCTCACTGGCCAGTGTACAGCAA 250
  |||||
17 sArgLyLeuProCysAspThrProGlyGlnGlyLeuThrCysSerGly 34
  |||||
251 ACGCAATTGCCATGTGTAATCCAGAGCAAGGCTTAACCTGACAGTGTG 300
  |||||
34 LysArgArgArgGluGlnGluSerLysTyrIleGluGluLeuAlaGlu 50
  |||||
301 AAAAAGCGAGACGGAGACGAAAGTAATATATATGAAGATGGCTGAG 350
  |||||
51 LeuIleSerAlaAsnLeuSerAspIleAspAsnPheAsnValLysProAs 67
  |||||
351 CTGATATCTGCCATCTTAGTATGATGACAAATTCATATGCAACAGCA 400
  |||||
67 pLysCysAlaIleLeuLysGluThrValArgGlnIleArgGlnIleLysG 84
  |||||
401 TAAATGTCGCAATTTAAAGAAACAGTAAGACATGATCGTCAATTAAGA 450
  |||||
84 LysGlnGlyLysThrIleSerAsnAspAspValGlnLysAlaAspVal 100
  |||||
451 AGCAGAGAAACCTATTTCCATATGATGATGATGATGATGATGATGATG 500
  |||||
101 SerSerThrGlyGlnGlyValIleAspLysAspSerLeuGlyProLeu 117
  |||||
501 TCTTCTACAGGGCAGGGAGTATGATTAAGACTCTTAGACCGCTTTT 550
  |||||
117 uLeuGlnAlaLeuAspGlyPheLeuPheValAlaAsnArgGlnAlaAsn 134
  |||||
551 ACTTCAGGCAATGGATGTTCTCTATTTGTGTAATGAGACGAGAAACA 600
  |||||
134 lValAlaPheValSerGluAsnValThrGlnTyrLeuGlnTyrLysGln 150
  |||||
601 TTGATTTGTATCAGAAATGTCACACAAATCTGCAATATATACAGAG 650
  |||||
151 AspleuValAsnThrSerValTyrAsnIleLeuHISGluGluAspArgly 167
  |||||
651 GACCTGGTTACACAGGTTTACAAATCTTACATGAGAGAGACAGAAA 700
  |||||
167 sAspPheLeuLysAsnLeuProLysSerThrValAsnGlyValSerTrp 184
  |||||
701 GGATTTTCTTACAGATTTTACCAAAATCTACAGTTAATGAGTTTCTG 750
  |||||
184 hAsnGluProGlnArgGlnLysSerHisThrPheAsnCysArgMetLeu 200
  |||||
751 CAAATGAGACCCAAAGACAAAGAACCATATTAATTTGCCGTATGTTG 800
  |||||
201 MetLysThrProHISAspIleLeuGluAspIleAsnAlaSerProGlu 217
  |||||
801 ATGAAACACCCACATGATTTCTGAGACATTAACGCCAGTCCGGAAT 850
  |||||
217 LArgLArgTyrGluThrMetGlnCysPheAlaLeuSerGlnProArg 234
  |||||
851 GCGCAGAGATATGAACAATGAGTGTGTCCTGCTCAGCAGCAGAG 900
  |||||
234 lMetMetGluGluGlyLysAsnLeuGlnSerCysMetIleCysValAla 250
  |||||
901 CTATGATGAGAGAGGAGGAAATTTGCAATCTTGTATGATCTGTGTGCA 950
  |||||
251 ArgArgIleThrThrGlnGlyLysArgThrPheProSerAsnProGluSerPh 267
  |||||
951 CGCCGATTAACACAGAGAAAGAAATTTCCATCAACACCTGAGACTT 1000
  |||||
267 eIleThrArgHisAspLeuSerGlyLysValValAlaAsnIleAspThrAsn 284
  |||||
1001 TATTACAGACATGATCTTTCAGGAAAGGTTGTCAAATATGATCAAAAT 1050
  |||||
284 eIleuArgSerSerMetArgProGlyPheGluAspIleIleArgArgCys 300
  |||||
1051 CACTGAGATCTCCATAGAGGCTGCTTTGAGATATATATCCAGAGTGT 1100
  |||||
301 lIleGlnArgPhePheSerLeuAsnAspGlyLysIleSerTrpSerGlnLys 317
  |||||
1101 ATTCAGAGATTTTATGCTAATATGATGGAGAGATGATGCTCCAGAAAC 1150
  |||||
317 gHisTyrGlnGluAlaTyrLeuAsnGlyHisAlaGluThrProValTyr 334
  |||||
1151 TCACATACAGAAAGCTTATCTTAATGGCCATGAGAAACCCAGTATATC 1200
  |||||
334 rPheSerLeuAlaAspGlyThrIleValThrAlaGlnThrLysSerLys 350
  |||||
1201 GATTCGTGTTGGCTGATGAACTATAGTACAGCAGACAGAAAGCAAA 1250
  |||||
351 LeuPheArgAsnProValThrAsnAspArgHisGlyPheValSerThrH 367
  |||||
1251 CTCTTCGAAATCTGTAAACAAATGATGACATGATGCTTTGTCTCAACCA 1300
  |||||
367 sPheLeuGlnArgGluGlnAsnGlyTyrArgProAsnProAsnProValG 384
  |||||
1301 CTTCCTTCAGAGAACAGATGATATGACCAACCCAAATCTGTG 1350
  |||||
384 LysGlnGlyIleArgProMetAlaGlyCysAsnSerSerValGlyGly 400
  |||||
1351 GACAAGGATTAAGACCACTATGCTGATGATCAACAGTGGTGGAGGGGC 1400
  |||||
401 MetSerMetSerProAsnGlnGlyLeuGlnMetProSerSerArgAlaTyr 417
  |||||
1401 ATGAGTATGTGCCAACAACAGGCTTACAGATGCGCAGACAGACAGGCT 1450
  |||||
417 rGlyLeuAlaAspProSerThrThrGlyGlnMetSerGlyValAlaArgTyrG 434
  |||||
1451 TGGCTTGACAGACCCCTAGACACACAGGAGATGATGAGACTAGTATG 1500
  |||||
434 LysGlySerSerAsnIleAlaSerLeuThrProGlyProGlyMetGlnSer 450
  |||||
1501 GGGGTTCCAGTAACATGACTTCAATGACCCCTGGGCCAGGATCAATCA 1550
  |||||
451 ProSerSerTyrGlnAsnAsnAsnTyrGlyLeuAsnMetSerSerProPr 467
  |||||
1551 CCATCTTCTCCAGAACACACACTATGAGGCTCAACATGATGATACCCCC 1600
  |||||
467 oHisGlySerProGlyLeuAlaProAsnGlnGlnAsnIleMetIleSerP 484
  |||||
1601 ACATGGGAGGTCCTGCTGTCGCCCAACACAGACAGATATATGATTTCTC 1650
  |||||
484 rAlaGlnArgGlySerProLysIleAlaSerHisGlnPheSerProVal 500
  |||||
1651 CTCGTAACCGGGAGTCCAAAGATAGGCTCACAATGATTTTCTCCGTT 1700
  |||||
501 AlaGlyValHisSerProMetAlaSerSerGlyAsnThrGlyAsnHis 517
  |||||
1701 GCAGGTGACACTCTCCCATGATGCTTCTGCAATATGAGGAACACAGAG 1750
  |||||
517 rPheSerSerSerSerLeuSerAlaLeuGlnAlaIleSerGlnGlyValG 534
  |||||
1751 CTTTTCAGAGCTCTCTCACTGAGCCCTGCAAGCCATGAGTGAAGTGTG 1800
  |||||
534 LThrSerLeuLeuSerThrLeuSerSerProGlyProLysLeuAspAsn 550
  |||||
1801 GGACTTCCCTTTATCTACTCTGATCACCAGGAGCCCAATTTGGATAAC 1850
  |||||
551 SerProAsnMetAsnIleThrGlnProSerLysValSerAsnGlnAspSe 567
  |||||

```

1851 TCTCCCATATGATATTAACCAACCAAGTAAAGTAAAGCAATCAGAGTTC 1900  
567 rLysSerProLeuGlyPheTyrCysAspGlnAsnProValGluSerSerM 584  
1901 CAAGAGTCTCTGGGCTTTATTGGACCAAAATCCAGTGGAGAGTTCAA 1950  
584 eLcysGlnSerAsnSerArgAspHisLeuSerAspLysGluSerLysGlu 600  
1951 TGTGTGAGTCAAAATAGACAGATCACCTCCAGTACAAAGAAAGTAAAGAG 2000  
601 SerSerValGluGlyAlaGluAsnGlnArgGlyProLeuGluSerLysG 617  
2001 AGCAGTGTGGAGGGGCGACAGAAATCAAAAGGCTCTTGGAAAGCAAGG 2050  
617 yHisLysLysLeuLeuGlnLeuLeuThrCysSerSerAspAspArgGlyH 634  
2051 TCATATAAAATTAAGTACAGTACTTACCTGTTCTTGATGACCGGGGTC 2100  
634 lAsSerSerLeuThrAsnSerProLeuAspSerSerCysLysGluSerSer 650  
2101 ATTCCCTCCTTGACCAACTCCCTCCTAGATTCAAGTTGTAAGAAATCTTCT 2150  
651 ValSerValThrSerProSerGlyValSerSerSerThrSerGlyGlyA 667  
2151 GTTAACTGTCACCAAGCCCTCTGGAGTCTCTCTTACATCTGGAGAGGT 2200  
667 lSerSerThrSerAsnMetHisGlySerLeuLeuGlnGluLysHisArgI 684  
2201 ATCCCTCATCATCAATATGCATGGGTCACTGTACAAAGAAAGCAACCGGA 2250  
684 lLeuHisLysLeuLeuGlnAsnGlyAsnSerProAlaGluValAlaLys 700  
2251 TTTTGCACAAAGTTGCTGCAGAAATGGAAATTCACACAGCTGAGTAGCCAG 2300  
701 lIleThrAlaGlnAlaThrGlyLysAspThrSerSerIleThrSerCysG 717  
2301 ATTACTGCAGAAAGCCACTGGAAAGACACACAGATATACTTCTTGTCG 2350  
717 yAspGlyAsnValValLysGlnGluGlnLeuSerProLysLysGlyAla 734  
2351 GGACGGAAATGTTGTCCAGACAGACAGCTTAAGTCTTAAGAAAGAGACA 2400  
734 snAsnAlaLeuLeuArgTyrIleLeuLeuAspArgAspAspProSerAspAla 750  
2401 ATAAATGCACTTCTTAGATACCTGCTGGACAGGAGATCTTAGTGAATGA 2450  
751 LeuSerLysGluLeuGlnProGlnValGluGlyValAspAsnLysMetSe 767  
2451 CTCTCTAAAGAACTACAGCCCAAGTGGAAAGAGTGAATATAATGAG 2500  
767 rGlnCysThrSerSerThrIleProSerSerSerGlnGluLysAspProL 784  
2501 TCAGTGCACACCTCCACCATCTCTAGCTCAAGTCAAGCAAGAAAGCCCTTA 2550  
784 yAlaLysThrGluThrSerGluGluGlySerGlyAspLeuAspAsnLeu 800  
2551 AAATTAAGACAGAGACAGTGAAGAGGATCTGGAGACTTGGATATATCA 2600  
801 AspAlaIleLeuGlyAspLeuThrSerSerSerAspPheTyrAsnAsnSerI 817  
2601 GATGCTATTCTTGGTGATCTGACTAGTTCTGACTTTTACAATAATTCAT 2650  
817 eSerSerAsnGlySerHisLeuGlyThrLysGlnGlnValPheGlnGlyT 834  
2651 ATCCCTCAATGCTAGTCACTCTGGGACTTAAGCAACAGGTGTTCAAGGAA 2700  
834 hAsnSerLeuGlyLeuLysSerSerSerGlnSerValGlnSerIleArgPro 850  
2701 CTAAATCTCTGGGTTGAAAAGTTCAACACTGTGCAGCTATTCGTCTCT 2750  
851 ProTyrAsnArgAlaValSerLeuAspSerProValSerValGlySerSe 867  
2751 CCATATAACCGAGAGTGTCTGTGATAGCCCTGTTTCTGTGGCTCAAG 2800

867 rProProValLysAsnIleSerAlaPheProMetLeuProLysGlnProM 884  
2801 TCTCCAGTAAAAAATATCAGTGTCTTCCCAATGTTACCAAGAACCA 2850  
884 eLLeuGlyGlyLysProArgMetMetAspSerGlnGluAsnTyrGlySer 900  
2851 TGTGGGTGGGAATCCAGAAATGATGATGATGATGATGATGATGATGATGAT 2900  
901 SerMetGlyGlyProAsnArgAsnValThrValThrGlnThrProSerSe 917  
2901 AGTATGGGTGGGCCCAACCGAAATGTGACTGTACTCAGACTCTTCTCTC 2950  
917 rGlyAspThrGlyLeuProAsnSerLysAlaGlyArgMetGluProMetA 934  
2951 ACAGAGCTGGGCTTACCAAACTCAAAAGCCGCGAAGATGAACTTATGA 3000  
934 snSerAsnSerMetGlyArgProGlyGlyAspTyrAsnThrSerLeuPro 950  
3001 ATTCAAACTCCATGGGAAACAGACAGAGAGATTAATATATCTTTTACC 3050  
951 ArgProAlaLeuGlyGlySerIleProThrLeuProLeuArgSerAsnSe 967  
3051 AGACCTGCACCTGGGTGGCTCTATTCCACATTCCTCTTCCGTTAATAG 3100  
967 rIleProGlyAlaArgProValLeuGlnGlnGlnGlnMetLeuGlnM 984  
3101 CATACAGAGTGCAGACACAGTATTCACACAGACAGCAGATGCTTCAAA 3150  
984 eArgProGlyGlyLutLeuProMetGlyMetGlyAlaAsnProTyrGlyGln 1000  
3151 TGAGGCTGTGTAAATCCCATGGGAATGGGGCTTAATCCCTATGACCA 3200  
1001 AlaAlaAlaSerAsnGlnLeuGlySerThrProAspGlyMetLeuSerMe 1017  
3201 GCAGCAGCATCTAACCAACTGGGTTCTGCGCCGATGGCATGTTGCCAT 3250  
1017 rGluGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerL 1034  
3251 GGAAACAAGTTCTCATGGCACAACAAATAGGCCCTTCTTACGAATTTCCC 3300  
1034 euAspAspLeuValGlyProProSerAsnLeuGluGlyGlnSerAspGlu 1050  
3301 TCGATGATCTTTGTGGCCCACTTCCAACTCGAAAGCCAGATGACGAA 3350  
1051 ArgAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaTh 1067  
3351 AGAGCATTTATGGACACAGCTGCACACTTCTCAGCAACACAGATGCCAC 3400  
1067 rGlyLeuGlnGluIleAspArgAlaLeuGlyLutLeuProGluLeuValAsnG 1084  
3401 AGCCCTGGAAATTTGACAGAGCTTTGGGCATTCCTGAACTTGTCATTC 3450  
1084 lGlnGlnAlaLeuGluProLysGlnAspAlaPheGlnGlyGlnGluAla 1100  
3451 AGGGAACAGCATTAGAGCCCAACAGAGATCTTCCAAAGGCCAAGAAAGA 3500  
1101 AlaValMetMetAspGlnLysAlaGlyLeuTyrGlyGlnThrTyrProAl 1117  
3501 GCAGTAATATGATGATCAGAAAGCAGAGATTAATGAGACAGATACCAGC 3550  
1117 aGlnGlyProProMetGlnGlyGlyPheHisLeuGlnGlnGlnSerProS 1134  
3551 ACAGGGGCTCCCAATGCAAGAGGCTTTCATCTTCAGGAGCAATACCAT 3600  
1134 ePheAsnSerMetMetAsnGlnMetAsnGlnGlnGlyAsnPheProLeu 1150  
3601 CTTTAACTCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3650  
1151 GlnGlyMetHisProArgAlaAsnIleMetArgProArgThrAsnThrPr 1167  
3651 CAAGGAATCACCCACAGACCAACATCATGAGACCCCGGACAAACACCCC 3700

```

1167 oTysGlnLeuArgMetGlnLeuGlnGlnArgLeuGlnGlnGlnGlnPheL 1184
|||||
3701 CAGGCAACTAGATGAGCTTCAGCAGAGGCTGCAGAGGCGACAGCTTT 3750
|||||
1184 euAnGlnSerArgGlnAlaLeuGlnLeuLysMetGluAsnProThrAla 1200
|||||
3751 TGAATCAGAGCCGACAGGCACTGTAATTTGAAATGAAACCTACTGCT 3800
|||||
1201 GLYGlyAlaAlaAlaValMetArgProMetMetGlnProGlnGlnPheLe 1217
|||||
3801 GGtGGtGGtGGtGGtGATGAGGCTATCATCTACCCCGAGAGGCTTTTCT 3850
|||||
1217 uAnAlaGlnMetValAlaGlnArgSerArgGlnLeuLeuSerHisHisP 1234
|||||
3851 TAATGCTCAATGTGTCCTCCACAGCAGACAGAGCTCTAAGTCATCCTACT 3900
|||||
1234 hearGlnGlnArgValAlaMetMetMet.....GlnGln 1245
|||||
3901 TCCGACACACAGAGGCTGGCTATGATGATGATGATGATGATGATGATGAT 3950
|||||
1246 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1262
|||||
3951 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 4000
|||||
1262 ngInGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1279
|||||
4001 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 4050
|||||
1279 laSerProSerMetAspGlyLeuLeuAlaGlyProThrMetProGlnAla 1295
|||||
4051 CTTCCTCCACCATGATGATGGCTTTGGCAGAGCCACCAATGCCACAAGCT 4100
|||||
1296 ProProGlnGlnPheProTyrglnProAsnTyrglyMetGlyGlnGlnP 1312
|||||
4101 CCTCCGCAACAGTTTCCATATCAACCAAAATTATGATGGACACACACC 4150
|||||
1312 oaSPProAlaPheGlyArgValSerSerProProAsnAlaMetMetSers 1329
|||||
4151 AGATCCAGCCTTTGGTGGAGTGTCTAGTCTCTCCCAATGCATGATGTCGT 4200
|||||
1329 eaArgMetGlyProSerGlnAsnProMetMetGlnHisProGlnAlaAla 1345
|||||
4201 CAAGATAGGGTCCCTCCAGCAATCCATGATGCAACACCGGAGGTCGA 4250
|||||
1346 SerTleTyrglnSerSerGlnMetLysGlyTyrProSerSerLysAsnLeu 1362
|||||
4251 TCCATCTATCAGTCTCAGAAATGAAGGCTGCTCATCAGCAAGAAATTTGGC 4300
|||||
1362 aArgAsnSerSerPheSerGlnGlnGlnPheAlaHisGlnGlyAsnPro 1379
|||||
4301 CAGGACAGCTCCTTTCCAGCAGCAGAGTTGCCACACAGGAGGAAATCTCG 4350
|||||
1379 laValTyrgSerMetValHisMetAsnGlySerSerGlyHisMetGlyGln 1395
|||||
4351 CAGCTATATATATAGTGTGACATGATGAGCAGAGTGTGATCATGAGGAG 4400
|||||
1396 MetAsnMetAsnProMetProMetSerGlyMetProMetGlyProAspG 1412
|||||
4401 ATGAAATGAAACCCAGCCCATGCTGTGAGCATGCTATGGGTCCTGATCA 4450
|||||
1412 nLysTyrcys 1415
|||||
4451 GAATACTGC 4460

```

seq\_name: /cgn2.6/ptodata/1/pna/US6006\_COMB.seq:US-60-068-511-1

seq\_documentation\_block:

; Sequence 1, Application US/60068511

; GENERAL INFORMATION:

; APPLICANT: Suen, Chen-shian

; APPLICANT: Ftrail, Donald E.

; APPLICANT: Lytle, Richard C.

; TITLE OF INVENTION: Cloning and Expression of a Nuclear

```

; TITLE OF INVENTION: Receptor Coactivator Proteins and Uses Thereof
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Home Products Corporation
; STREET: One Campus Drive
; CITY: Parsippany
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/068-511
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Walsh, Andrea C.
; REGISTRATION NUMBER: 34,988
; REFERENCE/DOCKET NUMBER: 97243-00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 973-683-4117
; TELEFAX: 973-683-4117
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4789 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 185..4750
; US-60-068-511-1

alignment_scores:
  Quality: 7341.00      Length: 1423
  Ratio: 5.192          Gaps: 2
  Percent Similarity: 99.368  Percent Identity: 99.157

alignment_block:
  US-09-041-994-2 x US-60-068-511-1 ..

Align seg 1/1 to: US-60-068-511-1 from: 1 to: 4789

1 MetSerGlyLeuGlnGlnGlnLeuAspProLeuAlaSerAspSerArgly 17
|||||
185 ATGAGTGGATAGGAGAGAAACCTTGATCCAGTGGCCAGTGATTCACAGAA 234
|||||
17 sarGlyLsleuProCysAspThrProGlyGlnGlyLeuThrCysSerGly 34
|||||
235 ACCGCAATTCCTCAGTATCTCCAGACAAAGGCTTCACTGCAAGTGCTG 284
|||||
34 LuLysArgArgArgGlnGlnGlnSerLysTyrIleGlnGlnAlaGln 50
|||||
285 AAAAAGCGAGACGGAGCAGCAAGAACTAAATATATTTGAAGATTGGCTGAG 334
|||||
51 LeuIleSerAlaAsnLeuSerAspIleAspAsnPheAsnValLysProAs 67
|||||
335 CTGATATCTGCCAATCTTACTGATATGACAAATTTCAATGTCAAACACGA 384
|||||
67 pLysCysAlaIleLeuLysGlnThrValAlaArgGlnIleArgGlnIleLys 84
|||||
385 TAATGTGCGATTTTAAAGGAAACAGTAAGACAGATACGTCCTCAATTAAG 434
|||||
84 LuGlnGlyLysThrIleSerAsnAspAspValGlnLysAlaAspVal 100

```

435 AGCAAGGAAAACTATTTCAATGATGATGATTCAAAAGCCGATGTA 484  
 101 SerSerThrGlyGlnGlyValIleAspLysAspSerLeuGlyProLeuLe 117  
 485 TCTTCTACAGGGCAGGAGTATTATTAAGACTCTTACGAGCCCTTTT 534  
 117 uLeuGlnAlaLeuAspGlyPheLeuPheValValAsnArgGluAlaAsnI 134  
 535 ACTTCAGGCATTGGATGGTTTCTTATTTGTGTGTAATCGAGACGGAAACA 584  
 134 leValPheValSerGluAsnValThrGlnTyLeuGlnTyLeuGlnGlu 150  
 585 TTGTATTTTGTATCAGAAAATGTCACACAAATCTGCAAATATATAGCAAG 634  
 151 AsPheValAsnThrSerValTyAsnIleLeuNHISGluGluAspArgLy 167  
 635 GACCTGGTTAACACAAAGTGTTCACAAATCTTACATGAAGAAGACAGAAA 684  
 167 sAsPheLeuLysAsnLeuProLysSerThrValAsnGlyValSerTrpT 184  
 685 GGATTTTCTTAAGAATTACCAAAATCTACAGTTATGAGATTCTCTGGA 734  
 184 hrAsnGluProGlnArgGlnLysSerHisThrPheAsnCysArgMetLeu 200  
 735 CAAATGACACCCAAAGACAAAAGCCATACATTATTTCCGCTATGTTG 784  
 201 MetLysThrProNHISAspIleLeuGluAspIleAsnAlaSerProGluMe 217  
 785 ATGAANAACACACATGATATTCTGAGAACATTAACGCCAGTCTGAAT 834  
 217 LArgGlnArgTyrgLutThrMetGlnCysPheAlaLeuSerGlnProAla 234  
 835 GCGCCAGAGATATGAACAATGCAATGCTTGTGCTGTCTCAGCCAGAG 884  
 234 lAmelMeGluGluGlyLysAspleuGlnSerCysMetIleCysValAla 250  
 885 CTATCATGAGGAAGGGAAGATTTCGAATCTGTATGATCTGTGTGCA 934  
 251 ArgArgIleThrThrGlyGluArgThrPheProSerAsnProGluSerPh 267  
 935 CGCCGATTTACTACAGAGAAAAGACATTTCCATCAAAACCTGAGAGCTT 984  
 267 eLleThrArgHisAspleuSerGlyLysValValAsnIleAspTrpAsn 284  
 985 TATTACCGAATGATGATCTTTACGAAAGGTTGTCAAATAGATACAAATT 1034  
 284 eLleuArgSerSerMetArgProGlyPheGluAspIleIleArgArgCys 300  
 1035 CACGTAGATCTCTCATGAGGCTGGCTTTGAAGATTAATCCGAAAGGTGT 1084  
 301 IleGlnArgPhePheSerLeuAsnAspGlyGlnSerTrpSerGlnLysAr 317  
 1085 ATTCAAGATTTTATTAGTCTAAATGATGGCAGTCATGTCGCCAAGAACG 1134  
 317 gHisTyrgLutAlaTyLeuAsnGlyHisAlaGluThrProValTyra 334  
 1135 TCACATATCAAAAGCTTATCTTAAGGCCATGCAAAAACCCACAGATATATC 1184  
 334 rPheSerLeuAlaAspGlyThrIleValIThrAlaGlnThrLysSerLys 350  
 1185 GATTCCTGCTTGGCTGATGGAATAGTACTGCCACAGACAAAAGCAAAA 1234  
 1235 CTCTTCGAAATCCGTAAACAATGATGACATGGCTTTGTCTCAACCCA 1284  
 351 LeuPheArgAsnProValIThrAsnAspArgHisGlyPheValSerThrHI 367  
 367 sPheLeuGlnArgGluGlnAsnGlyTyArgProAsnProAsnProValG 384  
 1285 CTTCCTTCAGAGAACGAATGATATAGACCAAAACCAATCTCTGTG 1334  
 384 LysGlnGlyIleArgProProMetAlaGlyCysAsnSerSerValGlyGly 400

1335 GACAAAGGATTAGACCACTATGCTGGATGCAACAGTTCGGTAGCGCGC 1384  
 401 MetSerMetSerProAsnGlnGlyLeuGlnMetProSerSerArgAlaTy 417  
 1385 ATGAGTATGTGCCCAACCAAGGCTTACAGATGCCGACAGACAGGCGCTA 1434  
 417 rGlyLeuAlaAspProSerThrThrGlyGlnMetSerGlyAlaArgTyrg 434  
 1435 TGGCTTGGCAGACCCCTACACACAGGCGACAGATGATGAGCTAGATGTG 1484  
 434 LysGlySerSerAsnIleAlaSerLeuThrThrProGlyProGlyMetGlnSer 450  
 1485 GGGTTCCAGTACATAGCTTCATTAAGCCCTGGGCGCAGGCTGCAATCA 1534  
 451 ProSerSerTyrgLlnAsnAsnAsnTyrgLysLeuAsnMetSerSerProP 467  
 1535 CCATCTCTCTACACAGACACAACTATGGGTCACATGAGTAGGCCCCC 1584  
 467 OHISGlySerProGlyLeuAlaProAsnGlnGlnAsnIleMetIleSerP 484  
 1585 ACATGGAGTCTCTGGTCTTGCCCCCAACACAGCAGAAATATCATGATTCTC 1634  
 484 roArgAsnArgGlySerProLysIleAlaSerHisGlnPheSerProVal 500  
 1635 CTGCTAATCGTGGAGTCCAAAGATAGCCTCACAATGATTTTCTCTGTT 1684  
 501 AlaGlyValHisSerProMetAlaSerSerGlyAsnThrGlyAsnHisase 517  
 1685 GCAGGTGTGCACTCTCCATGGCATCTTGTGGCAATACTGGGAACACAG 1734  
 517 rPheSerSerSerSerLeuSerAlaLeuGlnAlaIleSerGluGlyValG 534  
 1735 CTTTTCCAGACGCTCTCTCAGTGGCCCTCAAGCCATGAGTAAGTGTGG 1784  
 534 LysThrSerLeuLeuSerThrLeuSerSerProGlyProLysLeuAspAsn 550  
 1785 GGACTTCCCTTTTATCTACTCTGTCATACACAGGCCCCCAATTTGATTAAC 1834  
 551 SerProAsnMetAsnIleThrGlnProSerLysValSerAsnGlnAspSe 567  
 1835 TCTCCCAATATGAATATTAACCAACCAAGTAAGTAACCAATACAGATTC 1884  
 567 rLysSerProLeuGlyPheTyrgCysAspGlnAsnProValGluSerSerM 584  
 1885 CAAGAGTCTCTGGGCTTTTATTCGCAACCAAAATCCAGTGAAGATTTCAA 1934  
 584 eCysGlnSerAsnSerArgAspHisLeuSerAspLysGluSerLysGlu 600  
 1935 TGTGTCACTCAAAATGCGAGATTCACCTCACTGACACAAAGAAATAGGAG 1984  
 601 SerSerValGluGlyAlaGluAsnGlnArgGlyProLeuGluSerLysGI 617  
 1985 AGCACTGTTGAGGGGCGAGAGATCAAAAGGGTCTTTGGAAGCAAAAGG 2034  
 617 yHisLysLysLeuLeuGlnLeuLeuThrCysSerSerAspAspArgGlyH 634  
 2035 TCATTAATAAATTACTGCAGTACTTACTGTTCTTCATATGACCGGGGTC 2084  
 634 lSerSerLeuThrAsnSerProLeuAspSerSerCysLysGlnSerSer 650  
 2085 ATTCTCTTTACACCACTCCCCCTAGATTCAAGTTGTAAAGAAATCTTCT 2134  
 651 ValSerValIThrSerProSerGlyValSerSerSerThrSerGlyGlyVa 667  
 2135 GTTAGTGTACACAGCCCTCTGAGATGTCGTCCTACATCTGGAGGAGT 2184  
 667 lSerSerThrSerAsnMetHisGlySerLeuLeuGlnGluLysHisArgT 684  
 2185 ATCCCTATCATCCAAATATGCAATGGTCACTGTACAAAGAGAAGCACCGGA 2234  
 684 leLeuHisLysLeuLeuGlnAsnGlyAsnSerProAlaGluValAlaLys 700  
 2235 TTTTGCACAAAGTTGCTGCAGAAATGGAAATTCACCAAGCTGAGGTAGCCAA 2284

701 IleThrAlaGlnAlaThrGlyLysAspThrSerSerIleThrSerCysG1 717  
 |||||||.....  
 2285 ATTTACTGAGAGCCATGGGAAAGACACCAGATATACCTTCTTGCG 2334  
 717 yAspGlyAsnValValLysGlnGlnLeuSerProLysLysGlnA 734  
 |||||||.....  
 2335 GGAGGAAATGTTGTCAAGCAGAGACACTTAAGTCTTAAGAAAGAG 2384  
 734 snAsnAlaLeuLeuArgTyrLeuLeuAspArgAspProSerAspAla 750  
 |||||||.....  
 2385 ATAATGCACTTCTTATGATACCTGTCAGACAGGATGATCCTAGTATCA 2434  
 751 LeuSerLysGlnLeuGlnProGlnValGlnGlyValAlaAspAsnLysMeSe 767  
 |||||||.....  
 2435 CTCTCTAAGAACTACAGCCCAAGTGGAGAGAGGATATATAAATAG 2484  
 767 rGlnCysThrSerSerThrIleProSerSerSerGlnGlnLysAspProL 784  
 |||||||.....  
 2485 TCACTGGACCAAGCTCCACCATCTCTAGCTCAAGTCAAGAAAGACCTTA 2534  
 784 yS1IleLysThrGlnThrSerGlnGlySerGlyAspLeuAspAsnLeu 800  
 |||||||.....  
 2535 AAATTAAGACAGACAAAGTGAAGAGGATCTGGAGACTTGATTAATCTA 2584  
 801 AspAlaIleLeuGlyAspLeuThrSerSerAspPheTyrAsnAsnSerIle 817  
 |||||||.....  
 2585 GATGCTATCTTGCTGATCTGACTAGTCTGACTTTCATTAACAATATCCAT 2634  
 817 eSerSerAsnGlySerHisLeuGlyThrLysGlnGlnValPheGlnGlyT 834  
 |||||||.....  
 2635 ATCCCTCAATAGTAGTCACTGGGAGACTAAGCAACAGGTGTTCAAGAA 2684  
 834 hAsnSerLeuGlyLeuLysSerSerGlnSerValGlnSerIleArgPro 850  
 |||||||.....  
 2685 CTAATTCCTGCTGGTGTGAAAAGTTCACAGTCTGCACACTTATTCGTCCT 2734  
 851 ProTyrAsnArgAlaValSerLeuAspSerProValSerValGlySerSe 867  
 |||||||.....  
 2735 CCAATTAACCAAGCAGATGCTCTGATATGCCCCCTGTTCTTGCGCTCAG 2784  
 867 rProProValLysAsnIleSerAlaPheProMetLeuProLysGlnProM 884  
 |||||||.....  
 2785 TCCTCCAGTAAATAATATACAGTGTTCCTCCATGTATCAAGAACCA 2834  
 884 eTLeuGlyLysAsnProArgMetMetAspSerGlnGlnAsnTyrGlySer 900  
 |||||||.....  
 2835 TGTGGGTGGGAATCCAAAGATGATGATAGTACAGGAAATTAATGCGCTCA 2884  
 901 SerMetGlyGlyProAsnArgAsnValThrValThrGlnThrProSerSe 917  
 |||||||.....  
 2885 AGTATGGGTGGGCCAAACCGAAATGTGACTGTGACTCAGACTCCTTCTCTC 2934  
 917 rGlyAspTyrGlyLeuProAsnSerLysAlaGlyArgMetGlnProMetA 934  
 |||||||.....  
 2935 AGGAGACTGGGGCTTACCAAACTCAAGCGCGCAAGATGGAACCTATGA 2984  
 934 snSerAsnSerMetGlyArgProGlyGlyAspTyrAsnThrSerLeuPro 950  
 |||||||.....  
 2985 ATTCAAACTCCATGGGAGAGACAGAGAGATTTAATAATCTCTTATACC 3034  
 951 ArgProAlaLeuGlyGlySerIleProThrLeuProLeuArgSerAsnSe 967  
 |||||||.....  
 3035 AGACTGTGACTGGGTGCTATTCCTCCACGTTGCTCTTCGCTTAATAG 3084  
 967 rIleProGlyAlaArgProValLeuGlnGlnGlnGlnMetLeuGlnM 984  
 |||||||.....  
 3085 CATACAGAGTGGAGACAGATATGCAACAGACAGACAGATGCTTCAAA 3134  
 984 eTArgProGlyGlnIleProMetGlyMetGlyAlaAsnProTyrGlyGln 1000  
 |||||||.....  
 3135 TGAGGCGCTGGTGAATCCCAATGGGAATGGGGCTTAATCCTATGGCCAA 3184

1001 AlaAlaIAserAsnGlnLeuGlySerThrProAspGlyMetLeuSerMe 1017  
 |||||||.....  
 3185 GCAGCAGCATTAACCAACTGGGTCTGCGCCGCAATGGCATGTGTCCAT 3234  
 1017 tGlnGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerL 1034  
 |||||||.....  
 3235 GGAAACAATTTCTCATGCGACTCAAAATAGGCGCTCTTGTAGAAATGCC 3284  
 1034 euAspAspLeuValGlyProProSerAsnLeuGlyGlnSerAspGln 1050  
 |||||||.....  
 3285 TGGATGATCTTGTGGGCGACCTTCAACCTGGAAGGCCAAGAGACGAA 3334  
 1051 ArgAlaLeuLeuAspGlnLeuHisThrLeuSerAsnThrAspAlaTh 1067  
 |||||||.....  
 3335 AGAGCATTTATGGACCAAGCTGCACACTTCTCAAGCAACACAGATGCCAC 3384  
 1067 rGlyLeuGlnGlnIleAspArgAlaLeuGlyTyrProGlyLeuValAsnG 1084  
 |||||||.....  
 3385 AGGCGTGAAGAATAATGACAGAGCTTTGGCATTCCTGAATCTGTCAATC 3434  
 1084 lngGlyAlaLeuGlnProLysGlnAspAlaPheGlnGlyGlnAla 1100  
 |||||||.....  
 3435 AGGACAGCGCATTTAGAGCCCAAGAGATGCTTCAAGGCCCAAGAGACA 3484  
 1101 AlaValMetMetAspGlnLysAlaGlyLeuTyrGlyGlnThrTyrProAl 1117  
 |||||||.....  
 3485 GCAGTAATGATGATCAAGAGCAGGATTAATGAGACAGACATGCCACGC 3534  
 1117 aglngIlyProProMetGlnGlyGlyPheHisLeuGlnGlnLysGlnSerProS 1134  
 |||||||.....  
 3535 ACAGGGCGCTCCATATGCAAGAGAGCTTTCATCTTCAGGAGCAATACCAT 3584  
 1134 erPheAsnSerMetMetAsnGlnMetAsnGlnGlnGlnGlyAsnPheProLeu 1150  
 |||||||.....  
 3585 CTTTAACTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3634  
 1151 GlnGlyMetHisProArgAlaAsnIleMetLysProArgTyrAsnThrPr 1167  
 |||||||.....  
 3635 CAAGGAATGCCACCAAGCAGACCATCATGAGACCCCGGACCAACACCCC 3684  
 1167 OlyGlyLeuLeuArgMetGlnLeuGlnGlnArgLeuGlnGlnGlnPheL 1184  
 |||||||.....  
 3685 CAAGCAACTTGAATGACAGCTTCAAGCAGAGCTCAGAGGCCACAGATTTT 3734  
 1184 euAsnGlnSerArgGlnAlaLeuGlnLeuLysMetGlnAsnProThrAla 1200  
 |||||||.....  
 3735 TGAATCAGAGCGCAGACAGCAGCTGATTAATGAAATGGAATGAAACCTACTGCT 3784  
 1201 GlyGlyAlaAlaValMetArgProMetMetGlnProGln.....G1 1214  
 |||||||.....  
 3785 GGTGGTGTGCGGTGATGAGGCTATGATGACAGCCCGAGTGAGCTCCCA 3834  
 1214 nGlyPheLeuAsnAlaGlnMetValAlaGlnArgSerArgGlyLeuLeuLys 1231  
 |||||||.....  
 3835 GGGTTTCTTATGCTCAATGATGTCGCCCAAGCAGCAGAGAGCTGCTAA 3884  
 1231 erHisHisPheArgGlnGlnArgValAlaMetMetMet..... 1243  
 |||||||.....  
 3885 GTCATCACTTCCGACCAAGAGGGGCTATGATGATGATGATGATGATGATGAT 3934  
 1244 ...GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1259  
 |||||||.....  
 3935 CAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 3984  
 1259 nGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1276  
 |||||||.....  
 3985 ACAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 4034  
 1276 snValThrAlaSerProSerMetAspGlyLeuLeuAlaGlyProThrMet 1292  
 |||||||.....  
 4035 ATGTGACTGCTTCCCGCAGCATGATGGCTTTTGGCAGGACCCCAATG 4084  
 1293 ProGlnAlaProProGlnGlnPheProTyrGlnProAsnTyrGlyMetG1 1309

```
|||||
4085 CCACAGCTCTCTCCGACAGTTTCCATATCAACCAATTTATGGAAATGGG 4134
1309 yGlnGlnProAspProAlaPheGlyValSerProProAlaAla 1326
4135 ACAACAACAGATCAGCCTTTGGTGGAGTGTAGTCCCTCCCAATGCAA 4184
1326 eMeLeSerArGMeGlyProSerGlnAsnProMeLeGlnHisPro 1342
4185 TGAATGCGCAAGAAATGGGTCCCTCCAGAAATCCCATGATGCAACCCG 4234
1343 GlnAlaAlaSerIleTyrGlnSerSerGlnMetLeuGlyTyrProSerG 1359
4235 CAGGCTGCATCCATCTATCACTGCTCTCAAGAAATGAAGGCTGGCCATCAG 4284
1359 yAsnLeuAlaArgAsnSerSerPheSerGlnGlnPheAlaHisGln 1376
4285 AATTTGGCCAGGACAGCCTTTTCCGACGACAGTTTGGCCACACAG 4334
1376 yAsnProAlaValTyrSerMetValHisMetAsnGlySerSerGlnHis 1392
4335 GGAATCCTCAGTGTATGATGTGTGCACATGAATGACAGCACTGGTCA 4384
1393 MetGlyGlnMetAsnMetAsnProMetProMetSerGlyMetProMetG 1409
4385 ATGGGACAGATGAACATGAACCCCATGCCCATGTGTGGCATGCTATGG 4434
1409 yProAspGlnLysTyrCys 1415
4435 TCCTGATCAGAAATACTGC 4453

seq_name: /cgn2_6/ptodata/1/pna/US093_COMB.seq:us-09-359-922-3962

seq_documentation_block:
: Sequence 3962, Application US/09359922
: GENERAL INFORMATION:
:   APPLICANT: Hyseq, Inc.
:   TITLE OF INVENTION: NOVEL CONTFICS OBTAINED FROM VARIOUS CDNA
:   TITLE OF INVENTION: LIBRARIES
:   FILE REFERENCE: 20411-752CON1
:   CURRENT APPLICATION NUMBER: US/09/359,922
:   CURRENT FILING DATE: 1999-07-22
:   EARLIER APPLICATION NUMBER: US 09/205,155
:   EARLIER FILING DATE: 1998-12-03
:   NUMBER OF SEQ ID NOS: 13203
:   SOFTWARE: FASTSEQ for Windows Version 3.0
:   SEQ ID NO 3962
:   LENGTH: 6855
:   TYPE: DNA
:   ORGANISM: Homo sapiens
:   FEATURE:
:   NAME/KEY: misc_feature
:   LOCATION: (1)...(6855)
:   OTHER INFORMATION: n = A,T,C or G
US-09-359-922-3962

alignment_scores:
    Quality: 7340.50      Length: 1424
    Ratio: 5.191          Gaps: 2
    Percent Similarity: 99.298    Percent Identity: 99.087

alignment_block:
US-09-041-994-2 x US-09-359-922-3962 ..
Align seg 1/1 to: US-09-359-922-3962 from: 1 to: 6855

1 MeSerGlyLeuGlyGlnAsnLeuAspProLeuAlaSerAspSerArgly 17
|||||
201 ATGAGTGGATTGAGAGAAAACTTGATCCATGGCCAGTGGATTCCAGAAA 250
17 sArgLysLeuProCysAspThrProGlyGlnGlyLeuThrCysSerGlyG 34
|||||
```

```
251 AGCAAAATGGCCATGTGATTAATCTCCAGACAAAGTCTTACCTGCAAGTGTG 300
34 LuLysArgArgArgGlnGlnGlnSerLysTyrIleGlnGlnLeuAlaGln 50
301 AAAAAGCGAGACGGAGAGCAAAATTAATATATGATGAGAAATGGCTGAG 350
51 LeuIleSerAlaAsnLeuSerAspIleAspAsnPhaAsnValLysProAs 67
351 CTGATATCTGCCCAATCTTATGTATTTGACAAATTTTCAATGTCAAACCAA 400
67 pLysCysAlaIleLeuLysGlnThrValArgGlnIleArgGlnIleLysG 84
401 TAAATGTGGCATTTTAAAGAAACAGTAAGACAGATACCTCAATATTAAG 450
84 LuGlnGlyLysThrIleSerAsnAspAspAspValGlnLysAlaAspVal 100
451 AGCAAGAAAACTATTTCCATGATGATGATGATGTCAAAAGCCGATGTA 500
101 SerSerThrGlyGlnGlyValIleAspLysAspSerLeuGlyProLeuLe 117
501 TCTTTCAACAGGCGAGGAGTTATTTGATAAAGACTCCTTAGGACCCCTTT 550
117 uLeuGlnAlaLeuAspGlyPheLeuPheValValAsnArgGlnAlaAsnI 134
551 ACTTCAGGCATTGGATGGATGTTCTTATTTGGTGAATGACAGACGAAACA 600
134 LeValPheValSerGlnAsnValThrGlnTyrLeuGlnTyrLysGlnGln 150
601 TTGTATTTGATCAGAAATGTCAACAAATACCTGCATATATTAAGCAAG 650
151 AspLeuValAsnThrSerValTyrAsnIleLeuHisGlnGlnAspArgly 167
651 GACCTGTTAACACAAAGTGTTCATATCTTACATGAAGAAAGACAGAAA 700
167 sAspPheLeuLysAsnLeuProLysSerThrValAsnGlyValSerTyrP 184
701 GGATTTTCTTAAGAAATTTACCAAAATCTACAGTTAATGAGTTTCTGTGA 750
184 hAsnGlnProGlnArgGlnLysSerHisThrPheAsnCysArgMetLeu 200
751 CAATGAGACCCCAAGACAAAAGCCATATCAATTAATTTGCGGTATGTTG 800
201 MetLysThrProHisAspIleLeuGlnAspIleAsnAlaSerProGluMe 217
801 ATGAAAAACACACATGATATTCGGAAGACATTAACGCCAGCTCTGAAT 850
217 LArgGlnArgTyrGlnThrMetGlnCysPheAlaLeuSerGlnProArg 234
851 GCCCCAGACATATGAACAATGCAGTGTCTTGCCCTGTCTCAGCCAGAG 900
234 LaMetMetGlnGlnGlyGlnAspLeuGlnSerCysMetIleCysValAla 250
901 CTATGATGGAGGAAGGGAAGATTTCATATCTGTATGATCTGTGGCA 950
251 ArgArgIleThrThrGlyGlnArgThrPheProSerAsnProGluSerPh 267
951 CGCCGCAATTAACAAGAGAAAGAACATTTCCATCAACCCCTGAGAGCTT 1000
267 eIleThrArgHisAspLeuSerGlyLysValValAsnIleAspThrAsn 284
1001 TATTTACCAGACATGATCTTTCAGAGAAAGTTGTCATATATAGATACAAT 1050
284 eLeuArgSerSerMetArgProGlyPheGlnAspIleIleArgArgCys 300
1051 CACTGAGATCTCTCATGAGAGCCGTGCTTGAAGATATTAATCCGAAGTGT 1100
301 IleGlnArgPhePheSerLeuAsnAspGlyGlnSerThrSerGlnLysAr 317
1101 ATTCAAGATATTTTATGTCTAAATGATGGGCAGTCTCCACAGAAAG 1150
317 gHisTyrGlnGlnAlaTyrLeuAsnGlnHisAlaGlnThrProValTyrA 334
1151 TCACATATCAAGAAAGCTTATCTTATATGCCCATGACAGAAACCCAGTATAT 1200
```

334 rphSerLeuAlaaspGlyThrIleValThrAlaGlnThrLysSerLys 350  
 |||  
 1201 GATTCTGGTGGCTGATGGAACTATAGTACTGCACAGACAAAAGCAAA 1250  
 351 LeuPheArgAsnProValThrAsnAspArgHisGlyPheValSerThrHis 367  
 |||  
 1251 CTCTCCGAATCTCTGTAACAATGATGCACATGGCTTTGTCTCAACCCA 1300  
 367 sPheLeuGlnArgGluGlnAsnGlyTyrArgProAsnProAsnProValG 384  
 |||  
 1301 CTTCCTTCAGAGAGAAACAGATGATATAGACCAACCAATCTGTGG 1350  
 384 lvgIneGlyIleArgProPromeTAlaGlyCysAsnSerSerValGlyGly 400  
 |||  
 1351 GACAAAGGATTTAGACCACTGATGGTGATGACAAAGTTCTGGTAGCGCG 1400  
 401 MetSerMetSerProAsnGlnGlyLeuGlnMetProSerSerArgAlaTyr 417  
 |||  
 1401 ATGATGATGTGGCCCAACCAAGGCTTACAGATGCCAGACGAGGCGCTTA 1450  
 417 rglYleuAlaaspProSerThrThrGlyGlnMetSerGlyAlaArgTyrG 434  
 |||  
 1451 TGGCTTGGCAGACCTAGCACACAGAGGCGAGATGAGTAGGATATG 1500  
 434 lvgLysSerAsnIleAlaSerLeuThrProGlyProGlyMetGlnSer 450  
 |||  
 1501 GGGGTTCCAGTAACATAGCTTCATTAAGCCCTGGCCAGGCAATGCAATCA 1550  
 451 ProSerSerTyrGlnAsnAsnAsnTyrGlyLeuAsnMetSerSerProP 467  
 |||  
 1551 CCATCTTCTACACAGAAACAACACTATGGGCTCAACATGATGATGCCCCC 1600  
 467 ohISGlySerProGlyLeuAlaProAsnGlnAsnIleMetIleSerP 484  
 |||  
 1601 ACATGGGAGCTCCGTCTTGCCCCAAACGACAGAAATATCATGATTTCTC 1650  
 484 roArgAsnArgGlySerProLysIleAlaSerHisGlnPheSerProVal 500  
 |||  
 1651 CTCTGATATGTGGAGTCCAAAGATAGCCCTCACATCATGTTTCTCCGTG 1700  
 501 AlaGlyValHisSerProMetAlaSerSerGlyAsnThrGlyAsnHisSe 517  
 |||  
 1701 GCAGGTGTGCACCTCCCATGGCATCTTTCGCAATACTGGGAACCAACAG 1750  
 517 rPheSerSerSerSerLeuSerAlaLeuGlnAlaIleSerGlyGlyValG 534  
 |||  
 1751 CTTTTCACAGAGCTCTCTCAGTGCCTCGCAAGCCATCAGTGAAGGTGTGG 1800  
 534 lYThrSerLeuSerThrLeuSerSerProGlyProLysIleAspAsn 550  
 |||  
 1801 GGACTTCCCTTTATCTACTCTCTCATCACAGGCCCAATGATGATTAAC 1850  
 551 SerProAsnMetAsnIleThrGlnProSerLysValSerAsnGlnAspSe 567  
 |||  
 1851 TCTCCCAATATGAAATTTACCACCAACAGTAAGTAAGCAATAGGATTC 1900  
 567 rLysSerProLeuGlyPheTyrCysAspGlnAsnProValGluSerSerM 584  
 |||  
 1901 CAAGAGTCTCTGTGGCTTTATATGCGACCAAAATCCAGTGCAGATGTCAA 1950  
 584 eTcysGlnSerAsnSerArgAspHisLeuSerAspLysGluSerLysGlu 600  
 |||  
 1951 TGTGTAGTCAAAATAGCAGATCACCTCAGTACAAAGAAATGAAGGAG 2000  
 601 SerSerValGluGlyValGluAsnGlnArgGlyProLeuGlnSerLysG 617  
 |||  
 2001 AGCAGGTGTAGGGGCAAGAAATCAAGGGTCTTTGGAAAGCAAAGG 2050  
 617 yHisLysLysLeuLeuGlnLeuLeuThrCysSerSerAspAspArgGlyH 634  
 |||  
 2051 TCATAAAAAATTTACTGACGTTACTACTCTGTCTGTGATGACCGGGGTC 2100

634 lSserSerLeuThrAsnSerProLeuAspSerSerCysLysGlnSerSer 650  
 |||  
 2101 ATTCTCTCTTGACCAACTCCCCCTAGATTCAAGTTGTAAGATCTTCT 2150  
 651 ValSerValThrSerProSerGlyValSerSerSerThrSerGlyGlyVal 667  
 |||  
 2151 GTTAGTGTACCAAGCCCTCTGTGAGTCTCTCTCTACATCTGTGAGAGAG 2200  
 667 lSerSerThrSerAsnMetHisGlySerLeuLeuGlnLysHisArgI 684  
 |||  
 2201 ATCCTCTACATCCAAATATGATGATGGTGCACCTGTACAGAGAAACACCGGA 2250  
 684 lLeuHisLysLeuLeuGlnAsnGlyAsnSerProAlaGluValAlaLys 700  
 |||  
 2251 TTTTGACAAAGTTGTGCAGAAATGGGAATTCACACAGCTGAGTAGCCAG 2300  
 701 lIeThrAlaGlnAlaThrGlyLysAspThrSerSerIleThrSerCysG 717  
 |||  
 2301 ATTACTGCAGAAACCACTGGGAAAGAACACAGCATATACTTCTGTGG 2350  
 717 yAspGlyAsnValValLysGlnGluLeuSerProLysLysLysGluA 734  
 |||  
 2351 GGACGGAAATGTTGTCAAGCAGAGCAGCTAACTCCTAAGAAAGAGAGAG 2400  
 734 sPAsnAlaLeuLeuArgTyrThrLeuLeuAspArgAspProSerAspAla 750  
 |||  
 2401 ATATGCACTTCTTAGATATACCTCTGTGACAGGGATATCTTAATGATGCA 2450  
 751 LeuSerLysGluLeuGlnProGlnValGluGlyValAlaAspAsnLysMetSe 767  
 |||  
 2451 CTCTCTTAAGAACTACACAGCCCAAGTGAAGAGTGGATATTAATTAATGAG 2500  
 767 rGlnCysThrSerSerThrIleProSerSerSerGlnGlyLysAspProI 784  
 |||  
 2501 TCAGTGCACAGCTCCACCACTTCCTAGCTCAAGTCAAGAAAGAACCCCTA 2550  
 784 ySileLysThrGluThrSerSerGluGluGlySerGlyAspLeuAspAsnLeu 800  
 |||  
 2551 AAATTTAAGACAGACAAAGTGAAGAGGATCTTGAGACTTGGATTAATCTTA 2600  
 801 AspAlaIleLeuGlyAspLeuThrSerSerAspPheTyrAsnAsnSerI 817  
 |||  
 2601 GATGCTATTCTGTGGTGTGATCTGACTAGTTCGACTTTACAAATTAATCCAT 2650  
 817 eSerSerAsnGlySerHisLeuGlyThrLysGlnGlnValPheGlnGlyT 834  
 |||  
 2651 ATCTCTCAAAAGTATGATCTGTGGGACTAAGCAACAGGTGTTTCAAGGAA 2700  
 834 hrasnSerLeuGlyLeuLysSerSerGlnSerValGlnSerIleArgPro 850  
 |||  
 2701 CTAAATCTCTGGGTTGAAAAGTTCAAGTCTGTGACGTTCTATTCGTCCT 2750  
 851 ProTyrAsnArgAlaValSerLeuAspSerProValSerValGlySerSe 867  
 |||  
 2751 CCATATAACCGACAGACGTCTCTGGAATAGCCCTGTTCTGTGTGGCTCAAG 2800  
 867 rProProValLysAsnIleSerAlaPheProMetLeuProLysGlnProm 884  
 |||  
 2801 TCTCTCAGTAAAAAATATCAGTCTTCCCATGTTACCAAAACCAACCA 2850  
 884 eTLeuGlyGlyAsnProArgMetMetAspSerGlnGluAsnTyrGlySer 900  
 |||  
 2851 TGTGGGTGGGAATCCAAAGATATGATGATGTCAGAAAAATTTATGGCTCA 2900  
 901 SerMetGlyGlyProAsnArgAsnValThrValThrGlnThrProSerSe 917  
 |||  
 2901 AGTATGGGTGGCCCAACCAAAATGTGACTGTGACTACACAGTCTCTCTC 2950  
 917 rGlyAspTrpGlyLeuProAsnSerLysAlaGlyArgMetGluPrometa 934  
 |||  
 2951 AGGAGACTGGGGCTTACCAAACTCAAAAGCGCGCAAGAAATGGAACCTATGA 3000  
 934 snSerAsnSerMetGlyArgProGlyGlyAspTyrAsnThrSerLeuPro 950

```

|||||
3001 ATTGAAATCCATGGAGACGACGAGAGATTAATATCTTCTTACC 3050
951 ArpProalaleuGlySerIleProthrleuProleuArgSerAsnSe 967
|||||
3051 AGACCTGCACCTGGGTGGCTTATTCACCATGCTCTTCGGTCTAATAG 3100
967 rIleProGlyAlaArgProValLeuGlnGlnGlnGlnImeLeuGlnm 984
|||||
3101 CATACAGAGTCGAGACCATGATTGCAACAGACAGACAGATGCTTCAAA 3150
984 etaArgProGlyGluIleProMetGlyMetGlyAlaAsnProTyrGlyGln 1000
|||||
3151 TGAGGCTGTGTAATCCCATGGGAATGGGGCTTAATCCCTATGGCCAA 3200
1001 ALAAlaAlaSerAsnGlnLeuGlySerTyrProAspGlyMetLeuSerMe 1017
|||||
3201 GCACGACGATCTTAACCACTGGGTTCTCGGCCCATGGCATGTGTCCAT 3250
1017 tGluGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerL 1034
|||||
3251 GGAACAAAGTTTCTCATGGCACTCAAAATAGGCTCTCTTGTAGAAATTCCC 3300
1034 euAspAspLeuValGlyProProSerAsnLeuGlnGlyGlnSerAspGlu 1050
|||||
3301 TGGATGATCTTGTGGGCCACCTTCCAACTGGAAGGCCAGAGTACGAA 3350
1051 ArgAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaTh 1067
|||||
3351 AGAGATTAATTGGACACGACCTGACACTCTTCTCAGCAACACAGATGCCAC 3400
1067 rGlyLeuGlnGluIleAspArgAlaLeuGlyTyrLeuProGluLeuValAsnG 1084
|||||
3401 AGGCTGTGAAATAATTGACAGAGCTTGTGGCATCTCTGAACTGTGCAATC 3450
1084 lngGlyAlaLeuGlnIleProLysGlnAspAlaPheGlnGlyGlnIle 1100
|||||
3451 AGGCAAGAGCTTAGAGCCCAACAGAGTGCTTTCAGAGGCCAGAACCA 3500
1101 AlaValMetMetAspGlnIleAlaGlyLeuTyrGlyGlnThrTyrProAl 1117
|||||
3501 GCAGTAATGATGATCAGACAGAGGATTAATGACAGACATACCACGAC 3550
1117 aGlnGlyProProMetGlnGlyGlyPheHisLeuGlnGlnGlnSerProS 1134
|||||
3551 ACAGGGGCTCCCAATGCAAGAGGCTTTCATCTTCAGGAAATCACCAT 3600
1134 erPheAsnSerMetMetAsnGlnImeTAsnGlnGlyAlaAsnPheProLeu 1150
|||||
3601 CTTTAACTCTATGATGATGATGATGAACCAAGCAAGCAATTTTCTCTC 3650
1151 GlnGlyMetHisProArgAlaAsnIleMetArgProArgThrAsnThrPr 1167
|||||
3651 CAAGGAATGACCCACGACCAACATCATGAGACCCCGACAAACACCCG 3700
1167 oLyGlnLeuArgMetGlnLeuGlnGlnArgLeuGlnGlnGlnPheL 1184
|||||
3701 CAAGCAACTTGAATGACGCTTCAGACAGAGGCTGCAAGGCCAGCGTTT 3750
1184 euAsnGlnSerArgGlnAlaLeuGlnLeuLysMetGluAsnProThrAla 1200
|||||
3751 TGAATCAGAGCGACAGGCACTGTAATTTGAAATGAAAACCTTACTCT 3800
1201 GLyGlyAlaAlaValMetArgProMetMetGlnPro.....G1 1213
|||||
3801 GGTGGTCTCGCGTGATGAGGCTATGATGACCCCGAGTGAGCTCCCA 3850
1213 nGlnGlyPheLeuAsnAlaGlnImeValAlaGlnArgSerArgGluLeuL 1230
|||||
3851 GCAGGGTTTCTTAATGCTCAAAATGTGGCCCAACGACGAGAGGCTGC 3900
1230 euSerHisHisPheArgGlnGlnArgValAlaImeMetMet..... 1243
|||||

```

```

3901 TAACTCATCTCTCCGACACAGAGGCTGCTATGATGATGCACAGCAG 3950
1244 .....GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1258
|||||
3951 CAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAACAA 4000
1258 nGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1275
|||||
4001 GCAACAGCAACAGCAGCAACAGCAGCAACAGCAACAGGCTTCAGCCACCTC 4050
1275 roAsnValThrAlaSerProSerMetAspGlyLeuLeuAlaGlyProThr 1291
|||||
4051 CTATGTACTGCTTCCCGACAGATGATGGGCTTTTGGCAGGACCCACA 4100
1292 MetProGlnAlaProProGlnGlnPheProTyrGlnProAsnTyrGlyMe 1308
|||||
4101 ATGCCACAGCTCTCCGCAACAGTTTCCATATCACCAAAATTATGGAT 4150
1308 tGlyGlnGlnProAspProAlaPheGlyArgValSerSerProProAsnA 1325
|||||
4151 GGGACAAACACACAGATCCAGCCTTGGTGCAGTGTCTGTCTCCCAATG 4200
1325 lAmetMetSerSerArgMetGlyProSerGlnAsnProMetMetGlnHis 1341
|||||
4201 CAATGATGTGTCAGAAATGGGTCCCTCCAGAAATCCCATGATGCACAC 4250
1342 ProGlnAlaAlaSerIleTyrGlnSerSerGlnMetLysGlyTyrProSe 1358
|||||
4251 CCGCAGGCTGCATCCATCTATCAGTCCCTCAGAAATGAAAGGCTGGCCATC 4300
1358 rGlyAsnLeuAlaArgAsnSerSerPheSerGlnGlnGlnPheAlaHisG 1375
|||||
4301 AGGAATTTGGCCAGAACACCTCTTTCCACAGCAGCAATTTGCCACC 4350
1375 lngGlyAsnProAlaValTyrSerMetValHisMetAsnGlySerSerGly 1391
|||||
4351 AGGGAAATCTCGCAGTGTATGATGTGTCACATGAATGCCAGCAGTGT 4400
1392 HisMetGlyGlnMetAsnMetAsnProMetProMetSerGlyMetProMe 1408
|||||
4401 CACATGGGACAGATGACATGACATGACCCCATGTCATGTGCGATGCCAT 4450
1408 tGlyProAspGlnLysTyrCys 1415
|||||
4451 GGGTCTGTGATCAAAATCTGTC 4472
seq_name: /cgn2_5/prodata/1/pna/US093_COMB.seq:us-09-359-922-3962
seq_documentation_block:
; Sequence 3962, Application US/09359922A
; GENERAL INFORMATION:
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Liu, Jin
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
; FILE REFERENCE: 20411-752CON1
; CURRENT APPLICATION NUMBER: US/09/359,922A
; CURRENT FILING DATE: 1999-07-22
; EARLIER APPLICATION NUMBER: US 09/205,155
; EARLIER FILING DATE: 1998-12-03
; EARLIER APPLICATION NUMBER: US 09/034,341
; EARLIER FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 13203
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3962
; LENGTH: 6855
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(6855)
; OTHER INFORMATION: n = A,T,C or G
US-09-359-922-3962

```



## alignment\_scores:

Quality: 7340.50 Length: 1424  
 Ratio: 5.191 Gaps: 2  
 Percent Similarity: 99.298 Percent Identity: 99.087

## alignment\_block:

US-09-041-994-2 x US-09-359-922-3962 ..

Align seg 1/1 to: US-09-359-922-3962 from: 1 to: 6855

```

1 MetSerGlyLeuGlyIuAsnLeuAspProLeuAlaSerAspSerArgly 17
  |||||
201 ATGAGTGATAGGAGAAACTTGATGCCACTGCCAGTGATTCACGAAA 250
  |||||
17 sArgGlyLeuProCysAspThrProGlyGlyLeuThrCysSerGly 34
  |||||
251 ACGCAATTCATGTGATCTCCAGGACAAGCTTACCTGACAGTGCTG 300
  |||||
34 IuysArgArgArgGluGluGluSerIysTyrIleGluGluLeuAlaGlu 50
  |||||
301 AAAACCGAGACGGAGACGAGAAAGTAATATATTAAGAAATGGCTGAG 350
  |||||
51 LeuIleSerAlaAsnLeuSerAspIleAspAsnPheAsnValIysProAs 67
  |||||
351 CTGATATCTGCACATCTTAGTGATATGACATTTCAATGTCACAAACGAG 400
  |||||
67 pIysCysAlaIleLeuIysGluThrValArgGluIleArgGluIleIysG 84
  |||||
401 TAAATGTGCATTTTAAAGAAACAGTAAGCAGATACGTCATAATTAAG 450
  |||||
84 IuGlnGlyIysThrIleSerAsnAspAspValGlnIuysAlaAspVal 100
  |||||
451 AGCAGAGAAATATTTCCATGATGATGATGTTCAAAAAAGCCGATGTA 500
  |||||
101 SerSerThrGlyGlnGlyValIleAspIysAspSerLeuGlyProLeu 117
  |||||
501 TCTTCTACAGGCGAGGAGTATGATTAAGACCTTAGACCGCTTTT 550
  |||||
117 uLeuGlnAlaLeuAspGlyPheLeuPheValIaAsnArgGluAlaAsnI 134
  |||||
551 ACTTCAGGCAATTGGATGTTCTTCTATTTGGTGAATCGACAGCGAACA 600
  |||||
134 IeValIlePheValSerGluAsnValThrGlnIyrLeuGlnIyrIysGlnIu 150
  |||||
601 TTGTATTGTATCAGAAATGTCCACACATACCTGCAATATATTAACCAAGAG 650
  |||||
151 AspleuValAsnThrSerValTyrAsnIleuHisGluGluAspArgly 167
  |||||
651 GACCTGGTTAAACACAGTGTTTACAAATATCTTACATCAAGACAGACAGAA 700
  |||||
167 sAspPheLeuIysAsnLeuProIysSerThrValIaAsnGlyValSerTrp 184
  |||||
701 GGATTTTCTTAAGAAATTTACCAAAATCTACAGTTAATGAGATTTCCTGGA 750
  |||||
184 hrasnGluProGlnArgGlnIysSerHisThrPheAsnCysArgMetLeu 200
  |||||
751 CAAATGAGACCCAAAGCAAAAGCCATCATTTAATTTGCCGTATGTTG 800
  |||||
201 MetIysThrProHisAspIleLeuGluAspIleAsnAlaSerProGlu 217
  |||||
801 ATGAAAACACACATGATATTTCTGGAAGACATTAACCCGAGTCTGAAAT 850
  |||||
217 IArgGlnArgTyrGluThrMetGlnCysPheAlaLeuSerGlnProArg 234
  |||||
851 GCGCCAGAGATATGAACAATGACAGTCTTGCCCTGCTCAGCAGCAGAG 900
  |||||
234 IamMetGluGluGlyIuAspLeuIuSerCysMetIleCysValAla 250
  |||||
901 CTATGATGAGGAGGAGGAGGATTTGCAATCTTGATGATCTGTGTGCA 950
  |||||
251 ArgArgIleThrThrGlyIuArgThrPheProSerAsnProGluSerPh 267
  |||||
951 CGCGCATTTACTACAGGAGAAACATTTCCATCAACCACTGAGACTT 1000
  |||||
267 eIleThrArgHisIleAspLeuSerGlyIysValIaAsnIleAspThrAsn 284
  |||||
1001 TATTACCAAGACATGATCTTTCAGGAAAGGTTGTCAAATATATCAATAT 1050
  |||||
284 eIleuArgSerSerMetArgProGlyPheGluAspIleIleArgArgCys 300
  |||||
1051 CACTGAGATCTCCATGAGGCTGCTTGAAGATATATATCCGAAGGTGT 1100
  |||||
301 IleGlnArgPhePheSerLeuAsnAspGlyGlnSerTrpSerGlnIysAr 317
  |||||
1101 ATTCAGAGATTTTATGCTTAATGATGGAGCAGTCATGCTCCAGAAAGC 1150
  |||||
317 gHisTyrGlnGluAlaTyrIleuAsnGlyHisAlaGluThrProValTyr 334
  |||||
1151 TCACATCAAGAAAGCTTATCTTAATGGCCATGCAGAAACCCAGTATATC 1200
  |||||
334 rPheSerLeuAlaAspGlyThrIleValIThrAlaGlnThrIysSerIys 350
  |||||
1201 GATTCCTGTTGGCTGATGAGAACTATAGTACTGCACAGACAAAAAGCAA 1250
  |||||
351 LeuPheArgAsnProValThrAsnAspArgHisGlyPheValSerThrH 367
  |||||
1251 CTCTCCGAATTCCTGTAACAATGATGCAGATGCTTTGCTCAACCCA 1300
  |||||
367 sPheLeuGlnArgGluGlnAsnGlyTyrArgProAsnProAsnProValG 384
  |||||
1301 CTTCCTTCAGAGACAGAAAGATGATATAGACCAACCAATCTCTGTG 1350
  |||||
384 IyGlnGlyIleArgProProMetAlaGlyCysAsnSerSerValGlyGly 400
  |||||
1351 GACAGGAGATTAGACCACTATGCTGTGATGCAACAGTTCGTGTGGCGC 1400
  |||||
401 MetSerMetSerProAsnGlnGlyLeuGlnMetProSerSerArgAlaTy 417
  |||||
1401 ATGAGTATGTCGCCAACAACCAAGCTTACAGATGCCGAGCAGACAGGCTTA 1450
  |||||
417 rGlyLeuAlaAspProSerThrThrGlyGlnMetSerGlyValaArgTyrG 434
  |||||
1451 TGGCTTGACAGACCCCTAGACACACAGGACAGATGATGAGACTGATG 1500
  |||||
434 IyGlySerSerAsnIleAlaSerLeuThrProGlyProGlyMetGlnSer 450
  |||||
1501 GGGGTTCCAGTAACATGATGATGACCTTGCGCCAGGAGATCAATCA 1550
  |||||
451 ProSerSerTyrGlnAsnAsnAsnTyrGlyLeuAsnMetSerSerProP 467
  |||||
1551 CCATCTTCCTACCAACCAACACATATGGCTCAACATGATGACTAGCCCCC 1600
  |||||
467 oHisGlySerProGlyLeuAlaProAsnGlnGlnAsnIleMetIleSerP 484
  |||||
1601 ACATGGGAGTCCGTGCTTGCCCAACAGCAGCAATATCATGATTTCTC 1650
  |||||
484 rArgAsnArgGlySerProIysIleAlaSerHisGlnPheSerProVal 500
  |||||
1651 CTCGTAATCGGGAGTCCAAAGATAGCTTCACATCACTGATTTCTCCGTT 1700
  |||||
501 AlaGlyValHisSerProMetAlaSerSerGlyAsnThrGlyAsnHis 517
  |||||
1701 GCAGGTGCACTCTCCATGGCATCTTCTGCAATATCTGGGAACCAACAG 1750
  |||||
517 rPheSerSerSerSerLeuSerAlaLeuGlnAlaIleSerGluIyValG 534
  |||||
1751 CTTTTCAGAGAGCTCTCTCAGTGCCCTGCAAGCATCAGTAAGTAAGTGTGG 1800
  |||||
534 IyThrSerLeuLeuSerThrLeuSerSerProGlyProIysLeuAspAsn 550
  |||||
1801 GGACTTCCTTTATCTACTCTGTCATCACCAGGCCCAATATGGAATAC 1850
  |||||
551 SerProAsnMetAsnIleThrGlnProSerIysValSerAsnGlnAspSe 567
  |||||

```

```

1851 TCTCCCATATATGATATTACCAACCAAGTAAAGTAAGCAATCAGGATTC 1900
567 rLysSerProLeuGlyPheTyCysAspGlnAsnProValGluSerSer 584
|||||
1901 CAAGAGTCTCTGGGGTTTATTGGACCAAAATCAGTGAGAGTTCA 1950
584 eCysGlnSerAsnSerArgAspHisLeuSerAspLysGluSerLysGlu 600
|||||
1951 TGTGTGACGCAAAATAGACAGATCACCTCAGTACAAAGAAAGTAAGGAG 2000
601 SerMetGlyGluGlyAlaGlnAsnGlnArgGlyProLeuGluSerLysGlu 617
|||||
2001 AGCAGTGTGGAGGGGACAGAGATCAAGGGGCTCTTTGGAAAGCAAGAG 2050
617 yHisLysLysLeuLeuGlnLeuLeuThrCysSerSerAspAspArgGly 634
|||||
2051 TCATTAATAAAATTAAGTACAGTACTTACTGTTCTTCTGTATGACCGGGGTC 2100
634 sAspSerLeuThrAsnSerProLeuAspSerSerCysLysGluSerSer 650
|||||
2101 ATTCTCTCTTGACCAACTCCCTCAGATTCAAGTTGTAAAGAACTTCT 2150
651 ValSerValThrSerProSerGlyValSerSerSerThrSerGlyGly 667
|||||
2151 GTTAGTGTCACCAAGCCCTCTGAGTCTCTCTCTACATCTGGAGAGCT 2200
667 sSerSerThrSerAsnMetHisGlySerLeuLeuGlnGluLysHisArg 684
|||||
2201 ATCCCTACATCCAAATATGCATGGGTCACTGTACAAAGAAAGCAACCGGA 2250
684 LeuLeuHisLysLeuLeuGlnAsnGlyAsnSerProAlaGluValAlaLys 700
|||||
2251 TTTTGCACAAAGTTGCTGCAAGATGGAAATTCACAGCTGAGGTAGCCAAAG 2300
701 IleThrAlaGlnAlaThrGlyLysAspThrSerSerIleThrSerCysGlu 717
|||||
2301 ATTACTGCGAAGACCACTGGGAAAGACACACAGATATACTTGTGTGG 2350
717 yAspGlyAsnValValLysGlnGluGlnLeuSerProLysLysGluAla 734
|||||
2351 GGACGGAATGTTGTCAAGCAGACAGCTAAGTCCCTAAGAAAGAGAGA 2400
734 snAsnAlaLeuLeuArgTyrlaLeuLeuAspArgAspAspSerAspAla 750
|||||
2401 ATAAAGCACTTCTAGATACCTGTGTGACAGGAGGATCTTAGTGATGGA 2450
751 LeuSerLysGluLeuGlnProGlnValGluGlyValAspAsnLysMetSer 767
|||||
2451 CTCTCTAAAGAACTACAGCCCAAGTGGAAAGAGATATAAATAATGAG 2500
767 rGlnCysThrSerSerThrIleProSerSerSerGlnGluLysAspProL 784
|||||
2501 TCAGTGCAACACCTCCACCATCTCTAGCTCAAGTCAAGAGAAAGACCTTA 2550
784 ysiIleLysThrGlnThrSerGlnGluLysSerGlyAspLeuAspAsnLeu 800
|||||
2551 AATATTAAGACAGAGACAGTGAAGAGGATCTGGAGACTTGGATTAATCTTA 2600
801 AspAlaIleLeuGlyAspLeuThrSerSerAspPheTyAsnAsnSerIle 817
|||||
2601 GATGCTATTCTTGATCTGATCTGACTTCTTCAATAATATTCCT 2650
817 eSerSerAsnGlySerHisLeuGlyThrLysGlnGlnValPheGlnGlyT 834
|||||
2651 ATCCCTCAATGCTAGTCAATCTGGGACTTAAGCAACAGGCTTTCAAGGAA 2700
834 hrAsnSerLeuGlyLeuLysSerSerGlnSerValGlnSerIleArgPro 850
|||||
2701 CTAAATCTCTGGGTTTGAAAGTTACAGTCTGTGACAGCTATTGCTGCT 2750
851 ProTyAsnArgAlaValSerLeuAspSerProValSerValGlySerSer 867
|||||
2751 CCATATAACCGAGAGTGTCTGTGATAGCCCTGTTTCTGTGGCTCAAG 2800

```

```

867 rProProValLysAsnIleSerAlaPheProMetLeuProLysGlnPro 884
|||||
2801 TCCTCCAGTAAAAAATATCAGTGTCTTCCCAATGTTACCAAGCAACCCA 2850
884 eLeuGlyLysAsnProArgMetMetAspSerGlnGluAsnTyrlGlySer 900
|||||
2851 TGTGGGTGGGAATCCAAAGATGATGATAGTCAAGAAATATATGCTTCA 2900
901 SerMetGlyGlyProAsnArgAsnValThrValThrGlnThrProSerSe 917
|||||
2901 AGTATGGGTGGGCCAAACCGAAATGTGACTGTGACTCAAGCTCTTCTTC 2950
917 rGlyAspTrpGlyLeuProAsnSerLysAlaGlyArgMetGluProMet 934
|||||
2951 AGGAGACTGGGGCTTACCAAACTCAAAAGCCGCGAGAAATGGAACGTATGA 3000
934 snSerAsnSerMetGlyArgProGlyGlyAspTyrlAsnThrSerLeuPro 950
|||||
3001 ATTCAAACTCCATGGGAAACCAAGAGAGATTAATATATCTTCTTACCC 3050
951 ArgProAlaLeuGlyGlySerIleProThrLeuProLeuArgSerAsnSe 967
|||||
3051 AGACCTGCCTGGGTGCTCTATTCCACATTTGCTCTTGGCTTAATAG 3100
967 rIleProGlyAlaArgProValLeuGlnGlnGlnGlnMetLeuGln 984
|||||
3101 CATACAGGTGGCAGACCAAGTATGCAACAGCAGCAGAGATGCTTCAAA 3150
984 eArgProGlyLysIleProMetGlyMetGlyAlaAsnProTyrlGlyGln 1000
|||||
3151 TGAGGCTGTGTGAATCCCATGGAGATGGGGCTTAATCCCTATGGCCA 3200
1001 AlaAlaAlaSerAsnGlnLeuGlySerTrpProAspGlyMetLeuSerMe 1017
|||||
3201 GCAGACGATCTTAACCAACTGGGTTCTGTGCGCATGATGTGTCCAT 3250
1017 rGluGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerL 1034
|||||
3251 GGAACAAGTTTCTCATGGCAGCTCAAAATAGGCCCTCTTACGAATTTCCC 3300
1034 euAspAspLeuValGlyProProSerAsnLeuGlnGlyGlnSerAspGlu 1050
|||||
3301 TGGATGATCTTGTGGGCCACTTCCAACTGTGAAGGCCAGAGTACGGA 3350
1051 ArgAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThr 1067
|||||
3351 AGAGCATTTTGGACACAGCTGCACACTCTTCTCAGCAACACAGATGCCAC 3400
1067 rGlyLeuGlnGluIleAspArgAlaLeuGlyTlleProGluLeuValAsnG 1084
|||||
3401 AGCCTGGAAAGAAATGACAGAGCTTTGGGCAATTCCTGAAATCTTGCAATC 3450
1084 lngGlnAlaLeuGlnProLysGlnAspAlaPheGlnGlyGlnGluAla 1100
|||||
3451 AGGGAACGCGATTAGAGCCCAACAGAGATCTTCCAAAGGCCAAAGAAAGA 3500
1101 AlaValMetMetAspGlnLysAlaGlyLeuTyrlGlyGlnThrTyrlProAl 1117
|||||
3501 GCAGTAATATGATGATCAGAGCAGAGATTAATATGGACAGACATACCAGC 3550
1117 aGlnGlyProProMetGlnGlyGlyPheHisLeuGlnGlnGlnGlnSerProS 1134
|||||
3551 ACAGGGGCTCCCAATGCAAGAGAGCTTTCATCTTCAGGAGCAATACACAT 3600
1134 ePheAsnSerMetMetAsnGlnMetAsnGlnGlnGlyAsnPheProLeu 1150
|||||
3601 CTTTAACTCTATGATGATGATGATGATGATGATGATGATGATGATGATG 3650
1151 GlnGlyMetHisProArgAlaAsnIleMetArgProArgTyrlAsnThrPr 1167
|||||
3651 CAAGGAATGCAACCCAGAGCCAAACATCATAGACCCCGGACAAACACCCC 3700

```

```

1167 olvsglnleuargmetglnleuglnglnaargleuglnglyglnghpnel 1184
|||||
3701 CAGCAACTTAGATGACCTTCAGACAGGCGTCAGAGGCGCACACTTTT 3750
|||||
1184 euasnglnserargglnalauleglnleulysmetc1uasnprrhrala 1200
|||||
3751 TGAATCAGAGCCGACAGCACTTGATTAATGAAAAACCTACTGCT 3800
|||||
1201 glyglyalalavalmetargprometmetglnpro.....GI 1213
|||||
3801 GGGTGGTGGCGGTGATGAGGCGCTAGATGACAGCCCGAGTGAGCTCCA 3850
|||||
1213 nglnglypheleuasnalaglmetvalalaglnaargserargglnleul 1230
|||||
3851 GCAGGGTTTTCTTAATGCTCAAAATGCTCCGCCACAGCAGAGAGCTGC 3900
|||||
1230 euserhishispheargnglnaargvalalametmetc..... 1243
|||||
3901 TAAATGATCAGCTCCGACAACAGAGGGTGGCTATGATGACAGCAGAG 3950
|||||
1244 .....GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1258
|||||
3951 CAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 4000
|||||
1258 nglnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1275
|||||
4001 GCACACGACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 4050
|||||
1275 roasnvalthr1aserprosermetaspolyleuleualaglyprothr 1291
|||||
4051 CTAAATGTCAGCTGCTCCCGACAGATGATGGGCTTTGGCAGAGACCACA 4100
|||||
1292 metproGlnAlaProProGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1308
|||||
4101 ATGCCCAAGCTCTCCGACACAGTTTCATATCAACCAAAATTTGAAAT 4150
|||||
1308 tGlyGlnGlnProAspProAlaPheGlyArgVal1SerSerProProAsn 1325
|||||
4151 GGGACACACACAGATCCAGCTTTGGTGGAGTGTAGTCTCCCTCCCAATG 4200
|||||
1325 1ametwetserserargmetglyproserglnasnPrometmetglnhis 1341
|||||
4201 CAATGATGTCGTCAAGAAATGGGTCCCTCCAGAAATCCCATGATGCAAC 4250
|||||
1342 proGlnAlaAlaSer1LetyrglnSerSerGlnmetlysglytrprose 1358
|||||
4251 CCGCAGGCTGCATCCATCTATCATCTCAGAAATGAAAGGCTGGCCATC 4300
|||||
1358 rGlyAsnleuAlaArgAsnSerSerPheSerGlnGlnGlnPheAlaHisg 1375
|||||
4301 AGGAATTTGGCCAGACAGACTCTTTCCCGACAGCAGTTGGCCACC 4350
|||||
1375 lnglyAsnProAlaVal1YrSerMetVal1HisMetasnGlySerSergly 1391
|||||
4351 AGGGGAAATCTGCAGTGTATAGTATGTCACATGAATGACAGCAGTGGT 4400
|||||
1392 HisMetGlyGlnMetasnMetasnPrometPrometSerglyMettrone 1408
|||||
4401 CACATGTGGACAGATGAACATGAACCCCATGCCCATGTGGCATGTGCTAT 4450
|||||
1408 tGlyProAspGlnlySTYrCys 1415
|||||
4451 GGGTCTGTATCAGAAATACTGC 4472
|||||
seq_name: /cgn2_6/ptodata/1/pna/US6017_COMB.seq:US-60-172-373-13551

```

```

: TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide Polym
: FILE REFERENCE: GX-0006 P
: CURRENT APPLICATION NUMBER: US/60/172,373
: CURRENT FILING DATE: 1999-12-16
: NUMBER OF SEQ ID NOS: 25,772
: SOFTWARE: PERL Program
: SEQ ID NO 13551
: LENGTH: 7956
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc.feature
: OTHER INFORMATION: Incyte ID No: 44586.3
US-60-172-373-13551

alignment_scores:
    Quality: 7324.50      Length: 1424
    Ratio: 5.184          Gaps: 3
    Percent Similarity: 99.228      Percent Identity: 99.017

alignment_block:
US-09-041-994-2 x US-60-172-373-13551 ..

Align seg 1/1 to: US-60-172-373-13551 from: 1 to: 7956

1 MetSerGlyLeuGlyGlnLeuAspProLeuAlaSerAspSerArgly 17
|||||
216 ATAGTGGATTTAGAGAAAACTTGATCCAGTGGCCAGATTCACAGAAA 265
|||||
17 sarGlyLeuProCysAspThrProGlyGlnGlyLeuThrCysSergly 34
|||||
266 ACCGAAATGGCCATGTGATCTCAGACAGAGGCTTACCTGCAAGTGG 315
|||||
34 1uvsarGArGArGArgGlnGlnGlnSerLysTr1LeGlnGlnLeuAlaGln 50
|||||
316 AAAAAGGAGAGCGGAGCAGGAAAGTAATATTTGAAGATTGGCTGAG 365
|||||
51 Leu11Ser11Asn1LeuSerAsp11LeaspAsnPhasnVal1ysProAs 67
|||||
366 CTGATATCTGCCCAATCTTAGTATTTGACAAATTTCAATGTCAAACCA 415
|||||
67 pLysCysAla11LeuLysGlnThrVal1ArgGln11LeArgGln11Le 84
|||||
416 TAAATGGGATTTTAAAGAAACAGTAAAGACAGATACGTCAAATTAAG 465
|||||
84 1uGlnGlyLysThr11eSerAsnAspAspAspVal1GlnLysAlaAspVal 100
|||||
466 AGCAAGGAAAACTATTTCCAATGATGATGATGTCAAAAGCCGATGTA 515
|||||
101 SerSerThrGlyGlnGlyVal11easPlysAspSerLeuGlyProLeuLe 117
|||||
516 TCTTCTACAGGGCAGGAGTATTGATTAAGACTCCTTAGGACCGCTTT 565
|||||
117 uleuGlnAlaLeuAspGlyPheLeuPheVal1AlaAsnArgGlnAlaAsn 134
|||||
566 ACTTCAGGATTTGATGGATGTTCTATTTGGTGAATCGAGAGCAACCA 615
|||||
134 1eValPheValSerGlnAsnVal1ThrGlnTyLeuGlnIntyLysGlnGln 150
|||||
616 TTGTATTTTGTATCAGAAAATGTCCACAAATATCTGCAATTAAGCAAG 665
|||||
151 AspLeuValAsnThrSerVal1YrAsn11LeuHisGlnGlnAspArgly 167
|||||
666 GACCTGTTTAACACAAAGTGTTCAAATATCTTACATGAGAAAGACAGAA 715
|||||
167 sasPheLeuLysAsnLeuProLysSerThrVal1AsnGlyValSerTrpT 184
|||||
716 GGAATTTCTTAAGAAATTTACCAAAATCTACAGTTAATGAGGATTTCT 765
|||||
184 hrAsnGlnProGlnArgGlnLysSerHisThrPheAsnCysArgMetLeu 200
|||||
766 CAAATAGAGCCCAAGACAAAAAGCCATATTAATTTGCCGTATGTTGG 815
|||||

```

```

seq_documentation_block:
: Sequence 13551, Application US/60172373
: GENERAL INFORMATION:
: APPLICANT: Morris, MacDonald
: APPLICANT: Lal, Preethi
: APPLICANT: Diep, Dinh
: TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using

```

201 MetIysThrProHisAspIleLeuGluAspIleAsnAlaSerProGluMe 217  
|||||  
816 ATGAAACACACATGATATCTCGAAGACATAAACGCCACTCCTGAAT 865  
217 LArgGlnArgTyrGlnThrMetGlnCysPheAlaLeuSerGlnProArg 234  
|||||  
866 GCGCCAGAGATATGAACAATGAGTGTGGCTGTCTGACGACGAG 915  
234 LameMetGluGluGlyGluAspLeuGlnSerCysMetIleCysValAla 250  
|||||  
916 CTATGATGAGGAGGAGGAAATTTGCATCTTGTATGATCTGTGTGCA 965  
251 ArgArgIleThrThrGlyGluArgThrPheProSerAsnProGluSerP 267  
966 CGCCGCAATTACTACAGGAGAAACAATTTCCATCAAAACCTGAGAGCTT 1015  
267 eIleThrArgHisAspLeuSerGlyLysValValAsnIleAspThrAsn 284  
|||||  
1016 TATTACCGACATGATCTTTCAGGAAAGTTGTCAATATAGATTCAAAAT 1065  
284 eIleuArgSerSerMetArgProGlyPheGluAspIleIleArgArgCys 300  
|||||  
1066 CACGAGATCCTCCATGAGGCGCTTGGAATATATATCCGAAGCTGT 1115  
301 IleGlnArgPhePheSerLeuAsnAspGlyGlnSerTrpSerGlnLysAr 317  
|||||  
1116 ATTCAGAGATTTTATGCTAAATGATGGGCGATGCTCCAGAAACG 1165  
317 gHisTyrGlnGluAlaTyrLeuAsnGlyHisAlaGluThrProValTyrA 334  
|||||  
1166 TCACATTACAGAACCTTATCTTAATGGCCATGCAAGAAACCCCATATATC 1215  
334 rghSerLeuAlaAspGlyThrIleValThrAlaGlnThrLysSerLys 350  
|||||  
1216 GATTCGTGGTGGCTGATGGAACCTATAGACTGCACAGACAAAGCAAA 1265  
351 LeuPheArgAsnProValThrAsnAspArgHisGlyPheValSerThrH 367  
|||||  
1266 CTCTTCGAAATCCTGTAAACAAATGATCGACATGGCTTGTCTCAACCA 1315  
367 sPheLeuGlnArgGluGlnAsnGlyTyrArgProAsnProAsnProValG 384  
|||||  
1316 CTTCCTTCAGAGAACAGAAATGATATAGACCAACCAATCCTGTGTG 1365  
384 LngGlnGlyIleArgProPromeIaGlyCysAsnSerSerValGlyGly 400  
|||||  
1366 GACAAGGAGATTAGACCACTATGGCTGGATGCACAGTTGGTGGGCGC 1415  
401 MetSerMetSerProAsnGlnGlyLeuGlnMetProSerSerArgAlaTy 417  
|||||  
1416 ATGAGTATGTGCGCAACCAAGGCTTACAGATGCCGAGCAGGCGCTTA 1465  
417 rGlyLeuAlaAspProSerThrThrGlyGlnMetSerGlyAlaArgTyr 434  
|||||  
1466 TGGCTTGGCAACCTTAGCCACACAGGCGAGATGAGTGAAGTATG 1515  
434 LngIysSerSerAsnIleAlaSerLeuThrProGlyProGlyMetGlnSer 450  
|||||  
1516 GGGGTTCAGTAACATAGCTTATGACCCCTGGGCGAGCATGCATCA 1565  
451 ProSerSerTyrGlnAsnAsnAsnTyrGlyLeuAsnMetSerSerProP 467  
|||||  
1566 CCATCTTCCTCAACAGAACCAACTATGGGCTCAACATGAGTACGCCCCC 1615  
467 oHisGlySerProGlyLeuAlaProAsnGlnGlnAsnIleMetIleSerP 484  
|||||  
1616 ACATGGGAGTCTGTGCTTGGCCCAACACAGCAAGATATCATGATTTCTC 1665  
484 roArgAsnArgGlySerProLysIleAlaSerHisGlnPheSerProVal 500  
|||||  
1666 CTGTAATCGTGGGAGTCCAAAGATAGCCTCACATCAGTTTTCCTGCTT 1715

501 AlaGlyValHisSerProMetAlaSerSerGlyAsnThrGlyAsnHisSe 517  
|||||  
1716 GCAGGTGGCACTCTCCATGCGCATCTTCTGGCAATACTGGGAACACAG 1765  
517 rPheSerSerSerLeuSerAlaLeuGlnAlaIleSerGluGlyValG 534  
|||||  
1766 CTTTCCAGCAGCTCTCTCAGTGGCTGCAAGCCATCAGTGAAGGTGG 1815  
534 LyrThrSerLeuLeuSerThrLeuSerSerProGlyProLysLeuAspAsn 550  
|||||  
1816 GGACTCTCCCTTTATCTATCACTGTGCATCACCAGGCCCAATTTGGATTA 1865  
551 SerProAsnMetAsnIleThrGlnProSerLysValSerAsnGlnAspSe 567  
1866 TCTCCCATATGAATATTAACCAACCAAGTAAGTAAGCAATTCAGGATTC 1915  
567 rLysSerProLeuGlyPheTyrCysAspGlnAsnProValGluSerSerM 584  
|||||  
1916 CAAAGTCTCTGGGCTTTTATTCGACCAAAATCCAGTGAAGATTCAA 1965  
584 eCysGlnSerAsnSerArgAspHisLeuSerAspLysGluSerLysGlu 600  
|||||  
1966 TGTGTCACTCAAAATAGCAGAGATCACCTCAGTGACAAAGAAAGTAAGAG 2015  
601 SerSerValGluGlyAlaGluAsnGlnArgGlyProLeuGluSerLysG 617  
|||||  
2016 AGCAGGTGTGAGGGGCGAGAAATCAAAAGGGTCTTGGAAACCAAAAG 2065  
617 yHisLysLysLeuLeuGlnLeuLeuThrCysSerSerAspAspArgGlyH 634  
|||||  
2066 TCATRAAAATTAATGCGATTACTTACTCTTCTTTCATATACCGGGGTC 2115  
634 lSerSerLeuThrAsnSerProLeuAspSerSerCysLysGluSerSer 650  
|||||  
2116 ATTCCTCCTTGACCAACTCCGCCCTAGATTCAGTGTGAAGAAATCTTCT 2165  
651 ValSerValThrSerProSerGlyValSerSerSerThrSerGlyGlyVa 667  
|||||  
2166 GTTAGTGCACACACCCCTCGAGGTCCCTCTCATATCTGAGAGGAGT 2215  
667 lSerSerThrSerAsnMetHisGlySerLeuLeuGlnGluLysHisArg 684  
|||||  
2216 ATCTCTATCATCCAAATATGCAATGGGTCACTGTTCACAGACAAACCGGA 2265  
684 lLeuHisLysLeuLeuGlnAsnGlyAsnSerProAlaGluValAlaLys 700  
|||||  
2266 TTTTGCACAAGTGTGTCGAGAATGGGAATTCACCAGCTGAGGTAGCCAG 2315  
701 lIleThrAlaGlnAlaThrGlyLysAspThrSerSerIleThrSerCysG 717  
|||||  
2316 ATTACTGCAGAAAGCCACTGGGAAAGACACACAGATATTAACCTTGTGG 2365  
717 yAspGlyAsnValValLysGlnGlnGlnLeuSerProLysLysLysGlu 734  
|||||  
2366 GGACGGAAATTTGTCAAGCGAGAGACACTTAAGTCTTAAGAAAGAGAGA 2415  
734 snAsnAlaLeuLeuArgTyrLeuLeuAspArgAspAspProSerAspAla 750  
|||||  
2416 ATTAATGCACCTCTTAGATACCTGCTGGACAGGATGATCTTAAGTAAGCA 2465  
751 LeuSerLysGluLeuGlnProGlnValGluGlyValAspAsnLysMetSe 767  
|||||  
2466 CTCTCTTAAGAACTACACGCCCAAGTGAAGATGGCAATTAATAAGAG 2515  
767 rGlnCysThrSerSerThrIleProSerSerSerGlnGluLysAspProL 784  
|||||  
2516 TCAGTGCACCGCTCCACCATTTCTAGCTCAAGTCAAGCAAGAAAGACCTA 2565  
784 ysiIeLysThrGluThrSerGluGlnGlySerGlyAspLeuAspAsnLeu 800  
|||||  
2566 AAATTAAGACAGAGCAAGTAAGAGGGGATCTGGAAGACTTGATTAATCTA 2615  
801 AspAlaIleLeuGlyAspLeuThrSerSerAspPheTyrAsnAsnSerII 817

```

|||||
2616 GATGCAATCTTGATCTGACTAGTCTGACTTTTACAAATATTCAT 2665
817 eSerSerasnGlySerHisLeuGlyThrIysGlnGlnValPheGlnGlyT 834
2666 ATCTCAAAATGAGTATCATCTGGGGAGTAAAGCAAGGTCTTCAAGGAA 2715
834 hrasnSerLeuGlyLeuLysSerSerGlnSerValGlnSerIleLeuPro 830
2716 CTAATTCCTGGGGTTGAAAAAGTTACAGCTGTGCAGTCTATTCGGCTT 2765
851 ProTyrAsnArgAlaValSerLeuAspSerProValSerValGlySer 867
2766 CCATATACCGAGCAGCTCTCTGATAGCCCTGTTTCTTGGCTCAAG 2815
867 rProProValLysAsnIleSerAlaPheProMetLeuProLysGlnPro 884
2816 TCCTCCAGTAAAAAATATCAGTCTTCCCATGTTACCAAGCAACCCA 2865
884 eLeuGlyGlyAsnProArgMetMetAspSerGlnGlnAsnTyrGlySer 900
2866 TGTGGGTGGGAATCCAGATATGATAGTCAAGGAAAAATTATGGCTCA 2915
901 SerMetGlyGlyProAsnArgAsnValThrValThrGlnThrProSer 917
2916 AGTATGGTGGGCCCAACCGAAATGTGACTGTGACTCAGACTCTTCTCT 2965
917 rGlyAspTyrGlyLeuProAsnSerLysAlaGlyArgMetGluProMet 934
2966 AGGAGACTGGGGCTTACCAAACTCAAGCGCGGAGATGGAACCTATGA 3015
934 snSerAsnSerMetGlyArgProGlyGlyAspTyrAsnThrSerLeuPro 950
3016 ATTCAACTCTCATGGAGACACAGAGAGATATATATCTTCTTATCC 3065
951 ArgProAlaLeuGlyLysSerIleProThrLeuProLeuArgSerAsn 967
3066 AGACGTGCACATGGGTGGCTCTATTCACACTGCTCTTGGCTATAG 3115
967 rIleProGlyAlaArgProValLeuGlnGlnGlnGlnMetLeuGlnM 984
3116 CAAACACAGGTGCAGACAGTATTGCACAGCAGCAGCATGCTTCAAA 3165
984 eArg.....ProGlyGluIleProMetGlyMetGlyAlaAsnPro 997
3166 TGACTCCCCAGGGCTGTGGAAATCCCCATGGGAATGGGGCTAATCCC 3215
998 TyrGlyGlnAlaAlaSerAsnGlnLeuGlySerTyrProAspGly 1014
3216 TATGGCCAAAGCAGCATCTAACAACACTGGTTCCTGGCCGATGGCAT 3265
1014 LeuSerMetGlyGlnValSerHisGlyThrGlnAsnArgProLeuLeu 1031
3266 GTTGTCTCATGGAAAGTTTCTCATGACACTCAAAATAGGCTCTTCTTA 3315
1031 rAsnSerLeuAspAspLeuValGlyProProSerAsnLeuGlnGlyGln 1047
3316 GGAATTCCTGGATGATCTTGTGGGCACCTTCCAACTGGGAAGGCCAG 3365
1048 SerAspGluArgAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsn 1064
3366 AGTGACGAAAGACATTTATGGACAGCTGCACACTCTTCTCAGCAACAC 3415
1064 rAspAlaThrGlyLeuGlnGlnValLeuAspArgAlaLeuGlyTleProGlu 1081
3416 AGATGCCACAGGCGCTGGAAGAAATGACAGCTTGGGATTCCTGGAAC 3465
1081 euValAsnGlnGlyGlnAlaLeuGlnProLysGlnAspAlaPheGlnGly 1097
3466 TTGTCAATCAGGACAGGCTTATAGAGCCCAACAGGATGCTTCCAAAGC 3515
1098 GlnGlnAlaAlaValMetMetAspGlnLysAlaGlyLeuTyrGlyGln 1114
|||||

```

```

3516 CAAGACGACAGCAATATGATGATCAGAACGACAGATTATATGACAGAC 3565
1114 rTyrProAlaGlnGlyProProMetGlnGlyGlyPheHisLeuGlnGly 1131
3566 ATACCCAGCAGACAGGGCTTCCAATGCAGAGAGGCTTTCATCTTCAGGAG 3615
1131 InSerProSerPheAsnSerMetMetAsnGlnMetAsnGlnGlnGlyAsn 1147
3616 AATCACCATCTTTTAACCTATATGATGATGATGAACCGCAAGGCAT 3665
1148 PheProLeuGlnGlyMetHisProArgAlaAsnIleMetArgProArg 1164
3666 TTTCCTCTCCAGGAATGCACCCAGCAGCAATCATGAGACCCCGGAC 3715
1164 rAsnThrProLysGlnLeuArgMetGlnLeuGlnGlnArgLeuGlnGly 1181
3716 AAACACCCCAAGCAACTTGAATGCAGCTTTCAGCAGAGGCTCAGGGCC 3765
1181 InGlnPheLeuAsnGlnSerArgGlnAlaLeuGlnLeuLysMetGluAsn 1197
3766 AGCAGTTTGTGAATCAGAGCGAGCGACTTGAATGAAATGAAAC 3815
1198 ProThrAlaGlyAlaAlaValMetArgProMetMetGlnPro..... 1212
3816 CTTACTGCTGTGTGCTGTGGGTGATGAGGCTATGATGACAGGCCACGT 3865
1213 .....GlnGlnGlyPheLeuAsnAlaGlnMetValAlaGlnArgSera 1227
3866 GACCTCCACAGCAGGCTTCTTAATGCTCAAAATGCGCCCAACGACGA 3915
1227 rGlyLeuLeuSerHisHisPheArgGlnGlnArgValAlaMetMetMet 1243
3916 GAGAGCTGCTAAGTCAATCACTTCCGACACAGAGGGGTATGATGATG 3965
1244 .....GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1258
3966 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 4015
1258 nGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnAlaPheSerPro 1275
4016 GCACACAGCAACACAGCAGCAGCAGCAGCAACCCAGGCTTACGCCACCT 4065
1275 roAsnValThrAlaSerProSerMetAspGlyLeuLeuAlaGlyProThr 1291
4066 CTATGATGCTGCTTCCCGACATGATGGCTTTTGGCAGGACCCACA 4115
1292 MetProGlnAlaProProGlnGlnPheProTyrGlnProAsnTyrGly 1308
4116 ATGCCAACACTCTCTCGCAACAGTTTCCATATCAACCAATATATGGAAT 4165
1308 rGlyGlnGlnProAspProAlaPheGlyArgValSerSerProProAsn 1325
4166 GGGACAAACACAGATCCAGCTTGTGTCAGTGTAGCTCCCAATG 4215
1325 IamMetSerSerArgMetGlyProSerGlnAsnProMetMetGlnHis 1341
4216 CAATGATGTGTCAAGATGGTCCCTCCAGATGCCAGATGCAACAC 4265
1342 ProGlnAlaAlaSerIleTyrGlnSerSerGlnMetLysGlyTyrPro 1358
4266 CCGCAGGCTCATCATCTATCAGTCTCTGGAATGAAGGCTGGGCATC 4315
1358 rGlyAsnLeuAlaArgAsnSerSerPheSerGlnGlnGlnPheAlaHisG 1375
4316 AGGAAATTTGGCCAGGAAACAGCTCTTTTCCGACGACATTTGGCCACC 4365
1375 InGlyAsnProAlaValTyrSerMetValHisMetAsnGlySerSerGly 1391
4366 AGGGAAATCTCTCAGTGTATGATGTGCATGATGACAGCAGAGGGT 4415
1392 HisMetGlyGlnMetAsnMetAsnProMetProMetSerGlyMetPro 1408
4416 CACATGGGACAGATGAACATGAACCCCATGCTCTGGCATGCTCAT 4465

```

1408 tglyProaspGlnIysTyrCys 1415  
|||||  
4466 GGCTCTGATCATGAAATACGTC 4487  
seq\_name: /cgn2\_6/pcdata/1/pna/US094\_COMB.seq:us-09-440-612-1

seq\_documentation\_block:  
; Sequence 1, Application US/09440612  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Lex M. Cowser  
; TITLE OF INVENTION: ANTISENSE MODULATION OF SRC-3 EXPRESSION  
; FILE REFERENCE: RTS-0042  
; CURRENT APPLICATION NUMBER: US/09/440,612  
; CURRENT FILING DATE: 1999-11-15  
; NUMBER OF SEQ ID NOS: 47  
; SEQ ID NO 1  
; LENGTH: 6754  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (184)..(4422)  
US-09-440-612-1

alignment\_scores:  
Quality: 7234.50 Length: 1427  
Ratio: 5.175 Gaps: 3  
Percent Similarity: 97.968 Percent Identity: 97.758

alignment\_block:  
US-09-041-994-2 x US-09-440-612-1 ..

Align seg 1/1 to: US-09-440-612-1 from: 1 to: 6754

1 MetserGlyLeuGlyGluAsnLeuAspProLeuAlaSerAspSerArgly 17  
|||||  
184 ATGAGTGATTAGAGAAACTTGATCCACTGGCAGTGCATTCGCGAAA 233  
17 SARGLysLeuProCysAspThrProGlyGlnGlyLeuThrCysSerGlyG 34  
|||||  
234 ACGCAAAATGCGCATGTGATCTCCAGGACAAAGCTTACTCTGCGAGTGTG 283  
34 LuLysArgArgArgGluGlnLysSerTyrTleGluGluLeuAlaGlu 50  
|||||  
284 AAAAAGCGAGACGGAGCAGGAAAGTAATATATTGAAGAAATGGCTGAG 333  
51 LeuIleSerAlaAsnLeuSerAspIleAspAsnPheAsnValLysProAs 67  
|||||  
334 CTGATATCTCGCAATCTTAGTATGTGACAATTCATGTCAAACCGA 383  
67 PLYGcysAlaIleLeuLysGluThrValArgGlnIleArgGlnIleLysG 84  
|||||  
384 TAAATGTCGCAATTTTAAAGAAACGTAAGACAGTACGTCAATAAAG 433  
84 LuGlnGlyLysThrIleSerAsnAspAspValGlnLysAlaAspVal 100  
|||||  
434 AGCAAGGAAAAAATATTCCAAATGATGATGATTCAAAAGCCGATGTA 483  
101 SerSerThrGlyGlnGlyValIleAspLysAspSerLeuGlyProLeu 117  
|||||  
484 TCTTCTACAGGGCAGGAGGATTATGATTAAGACCTCTTAGAGACCGCTTT 533  
117 uLeuGlnAlaLeuAspGlyPheLeuPheValValAsnArgGlnAlaAsn 134  
|||||  
534 ACTTCAGGCAATGGATGCTTCTCTATTGCTGTGATGATGAGACGGAACA 583  
134 LeValPheValSerGluAsnValThrGlnTyrLeuGlnTyrLysGlnGlu 150  
|||||  
584 TTGTATTTGTATCAGAAAAATGTACACAAATACGTCAATATTAAGCAAG 633

151 AsPLeuValAsnThrSerValTyrAsnIleLeuHISGluAspArgly 167  
|||||  
634 GACCTGGTTAAACACAAAGTGTAAATATCTTACATGAAGAAGACAGAAA 683  
167 sAsPLeuLysAsnLeuProLysSerThrValAsnGlyValSerTyrP 184  
|||||  
684 GGATTTTCTTAAGAAATTTACCAAAATCTACAGTAAATGAGACTTCCGGA 733  
184 hrasGlnProGlnArgGlnLysSerHISThrPheAsnCysArgMetLeu 200  
|||||  
734 CAAATGAAACCCAAAGACAAAAAGCCATACATTTAATGGCCGATGTTG 783  
201 MetLysThrProHISAspIleLeuGlnAspIleAsnAlaSerProGlu 217  
|||||  
784 ATGAAACACCAACATGATATCTGAAAGACATAAACCGCACTCCGAAAT 833  
217 LArgGlnArgTyrGluThrMetGlnCysPheAlaLeuSerGlnProArg 234  
|||||  
834 GCGCCAGAGATATGAACAAATGCAAGTGTGGCTTGGCTGTCTACGCCAG 883  
234 lAmMetGluGluGlyLysAspLeuGlnSerCysMetIleCysValAla 250  
|||||  
884 CTATGATGAGGAAGGCGAAGATTGCAATCTGTGATGATCTGTGCGCA 933  
251 ArgArgIleThrThrGlyGluArgThrPheProSerAsnProGluSerP 267  
|||||  
934 CGCCGCAATCTACAGAGGAAAGAAACATTCATCAAAACCTCGAGAGCTT 983  
267 eLlEthrArgHISAspLeuSerGlyLysValValAsnIleAspThrAsn 284  
|||||  
984 TATTACAGACATGATCTTTCAGGAAAGTGTCAATATTAATACAAAT 1033  
284 eLeuArgSerSerMetArgProGlyPheGluAspIleIleArgArgCys 300  
|||||  
1034 CACTGAGATCCTCCATGAGCGCTTGCTTGAAGATATATATCGAAGGTGT 1083  
301 IleGlnArgPhePheSerLeuAsnAspGlyGlnSerThrPserGlnLysAr 317  
|||||  
1084 ATTCAAGATTTTATAGTCTAAATGATGGCAGTCAATGTCGCGAAACG 1133  
317 gHIStyrGlnGlu.....AlaTyrL 324  
1134 TCACATATAAAGATTTACCACTGATGGATATTTTCCCAACAGCTTATTC 1183  
324 euAsnGlyHISAlaGluThrProValTyrArgPheSerLeuAlaAspGly 340  
|||||  
1184 TTAATGCCCATGACAGAAACCCAGATATCGATTCTGCTGGCTGATGGA 1233  
341 ThrIleValIThrAlaGlnThrLysSerLysLeuPheArgAsnProValTh 357  
|||||  
1234 ACTATAGTACTGCACAGACAAAAAGCAAACTCTTCGAAATVCCCTGTATAC 1283  
357 rAsnAspArgHISGlyPheValSerThrHISpHeuGlnArgGluGln 374  
|||||  
1284 AAATGATGACATGGCTTGTCTCAACCCACTTCTTCAGAGAAACAGA 1333  
374 snGlyTyrArgProAsnProAsnProValGlyGlnGlyIleArgProPro 390  
|||||  
1334 ATGATATATAGACCAAAACCAATCTCTGTTGGACAAAGGATTAAGACCACT 1383  
391 MetAlaGlyCysAsnSerSerValGlyLysMetSerMetSerProAsnG 407  
|||||  
1384 ATGGCTGGATGCAACAGTTCGTTAGGCGCGCATGATGATGTGCCAAACCA 1433  
407 nGlyLeuGlnMetProSerSerArgAlaTyrGlyLeuAlaAspProser 424  
|||||  
1434 AGGCTTAAGATGCCGACAGCAGGCGCTTATGCTTGCGACAAACCTTACA 1483  
424 hrrThrGlyGlnMetSerGlyAlaArgTyrGlyLysSerAsnIleAla 440  
|||||  
1484 CCACAGGCGCAATGATGAGTACGTAGGTATGGGGTTCACAGTAACATAGCT 1533  
441 SerLeuThrProGlyProGlyMetGlnSerProSerSerTyrGlnAsnAs 457

1534 TCATTGACCCCTGGCCAGCATGCAATCCATCTCTACAGAACAA 1583  
 457 nasnTyGlyLeuAsnMetSerSerProHisGlySerProGlyLeuA 474  
 1584 CAACTTAGCTACACATGATGAGTCCGCCCATGAGAGTCTGCTTG 1633  
 474 laProAsnGlnInAsnIleMetIleSerProArgAsnArgGlySerPro 490  
 1634 CCCCAGAACGACAGAAATATCATATTTCTCTGTAATCGTGGAGTCA 1683  
 491 LysIleAlaSerHisGlnPheSerProValAlaGlyValHisSerProme 507  
 1684 AAGATAGCCCTACATCAGTTTCTCTGCTGAGGTGTCACCTCCAT 1733  
 507 LaIaSerSerGlyAsnThrGlyAsnHisSerPheSerSerSerLeus 524  
 1734 GGCATCTTCTGGCAATACCTGGGAACACACCTTTCCAGCAGCTCTCA 1783  
 524 eraIaLeuGlnAlaIleSerGluGlyValGlyThrSerLeuLeuSerThr 540  
 1784 GTCCCTCGAAGCCATCATGGAAGGTGGGACTTCCCTTTATCTACT 1833  
 541 LeuSerSerProGlyProLysLeuAspAsnSerProAsnMetAsnIleTh 557  
 1834 CTGTCAATCAACGACGCCCAAAATGGAATCACTCCCAATATGAAATATAC 1883  
 557 rGlnProSerLysValSerAsnGlnAspSerLysSerProLeuGlyPheT 574  
 1884 CCACACCAAGTAAAGTAAGCAATCAGATTCACAGAGCTCTGGGCTTTT 1933  
 574 yICysAspGlnAsnProValGluSerSerMetCysGlnSerAsnSerArg 590  
 1934 ATTGCAACCAAAATCCAGTGGAGAGTTCATGTCTCAATCAATATGAGCA 1983  
 591 AspHisLeuSerAspLysGluSerLysGluSerSerValGluGlyAlaG 607  
 1984 GATCACCCTGAGTACAAAGAAAGTAAGAGAGCAGTGTGAGGGGGAGCA 2033  
 607 uAsnGlnAlaGlyProLeuGlnIleSerLysGlyHisLysLysLeuLeuGln 624  
 2034 GAATCAAGGGGCTCTTGGAAAGCAAGTCAATAAAATTAATCTAGTCACT 2083  
 624 euleuThrCysSerSerAspAspArgLysIleSerSerLeuThrAsnSer 640  
 2084 TACTTACCTGTTCTTCTGATGACCGGGGTCAITCTCTTGACCAACTCC 2133  
 641 ProLeuAspSerSerCysLysGluSerSerValSerValThrSerProse 657  
 2134 CCCCTAGATCAAGTGTGTAAGAATCTTCTGTAGTGTACACAGCCCTTC 2183  
 657 rGlyValSerSerSerThrSerGlyValSerSerThrSerAsnMetH 674  
 2184 TGGAGTCTCTCTCTACATCATCTGAGAGATCTCTACATCCAAATATGC 2233  
 674 fsgIleSerLeuLeuGlnGluLysHisArgTlleuHisLysLeuLeuGln 690  
 2234 ATGGGTCACTGTTACAGAGAGACACCGGATTTGCAAGATGCTGCAG 2283  
 691 AsnGlyAsnSerProAlaGluValAlaLysIleThrAlaGlnAlaThrG 707  
 2284 AATGGGAATTCACACACTGAGTACGCAAGATTAATCTACACAGCCACTGC 2333  
 707 yLysAspThrSerSerIleThrSerCysGlyAspGlyAsnValValLysG 724  
 2334 GAAAGACACACAGCAGTATACTCTGTGGGACGGAATGTGTCAACG 2383  
 724 InGlnGlnLeuSerProLysLysLysGlnAsnAsnAlaLeuLeuArgTyr 740  
 2384 AGGAGCAGCTTAATCTCTTAAGAAAGAGAGAAATATGCACTTTAGATAC 2433  
 741 LeuLeuAspArgAspAspProSerAspAlaLeuSerLysGluLeuGlnPr 757

2434 CTGCTGGACAGGATGATCTCTAGTATGACATCTCTAAAGACTATACAGCC 2483  
 757 OGlnValGluGlyValAspAsnLysMetSerGlnCysThrSerSerThrI 774  
 2484 CCAAGTGGAGAGAGTGTATATAAAGACTAGTGCACAGCTCCACCA 2533  
 774 leProSerSerSerGlnGluLysAspProLysIleLysThrGluThrSer 790  
 2534 TTCTTACTCAAGTCAAGAGAAAGACCTTAATAATTAAAGACAGAGACAGT 2583  
 791 GluGluGlySerGlyAspLeuAspAsnLeuAspAlaIleLeuGlyAspLe 807  
 2584 GAAGAGGAGATCTGAGACTTGGAATATCTAGATCTATTCTTGATATCT 2633  
 807 uThrSerSerAspPheTyrAsnAsnSerIleSerSerAsnGlySerHisL 824  
 2634 GACTAGTCTGACTTTACATATATTCATATCTCTCAAAATGATGATCATC 2683  
 824 euGlyThrLysGlnGlnValPheGlnGlyThrAsnSerLeuGlyLeuLys 840  
 2684 TGGGAGCTAAGCAACAGCTTTTCAGAGAACTAATCTCTGGGTTTGAA 2733  
 841 SerSerGlnSerValGlnSerIleArgProProTyrAsnArgAlaValSe 857  
 2734 AGTTCAAGTCTGGAGTCTATTCGTCCTCCATATACCGAGACAGTGTCT 2783  
 857 rLeuAspSerProValSerValGlySerSerProProValLysAsnIleS 874  
 2784 TCTGGATAGCCCTGTTCTGTGGCTCAAGTCTCCAGTAAAAAATATCA 2833  
 874 eraIaPheProMetLeuProLysGlnProMetLeuGlyLysAsnProArg 890  
 2834 GTGCTTTCCCATGTTTACCAAGCAACCCATGTTGGTGGGAATCCACGA 2883  
 891 MetMetAspSerGlnGluAsnTyGlySerSerMetGlyLysProAsnAr 907  
 2884 ATGATGATATGTACAGAAATTAATGGCTCAAGTATG..... 2919  
 907 gAsnValThrValThrGlnThrProSerSerGlyAspTyrGlyLeuProA 924  
 2920 .....GGAGACTGGGGCTTACCA 2938  
 924 snSerLysAlaGlyArgMetGluProMetAsnSerAsnSerMetGlyArg 940  
 2939 ACTCAAGGCCGGCAGATGAACTATGAAATTCAAATCTCAATGGGAGCA 2988  
 941 ProGlyLysAspTyrAsnThrSerLeuProArgProAlaLeuGlyLysE 957  
 2989 CCAAGAGAGATTAATATCTTCTTATCCAGACCTGCACTGGGTGGCTC 3038  
 957 rIleProThrLeuProLeuArgSerAsnSerIleProGlyAlaArgProV 974  
 3039 TATTTCCCATTCCTCTCTGCTTAATATACATACACAGGTGCAGACAGC 3088  
 974 alleuGlnGlnGlnGlnMetLeuGlnMetArgProGlyLysIlePro 990  
 3089 TATTTCAACAGCAGCAGACAGATGCTTCAATAGAGCCTGGTGAATCC 3138  
 991 MetGlyMetGlyAlaAsnProTyrGlyGlnAlaAlaIleAsnAsnGlnLe 1007  
 3139 ATGGGAATGGGGCTATATCCCTATGGCCAGCAGCACTATACCACT 3188  
 1007 uGlySerTyrProAspGlyMetLeuSerMetGluGlnValSerHisGlyT 1024  
 3189 GGGTTCCTGGCCCGATGGCATGTTGTCTCATGAGAAAGATTTCTCATGGCA 3238  
 1024 hrGlnAsnArgProLeuLeuArgAsnSerLeusAspIleuValGlyPro 1040  
 3239 CTCAAAAATAGCCCTCTTCTTAGAATTTCCCTGGATGATCTTGTGGGCCA 3288  
 1041 ProSerAsnLeuGluGlnSerAspGluArgAlaLeuLeuAspGlnLe 1057  
 3289 CTTTCAACTGAGAGCCGAGAGTGAAGAAAGACATTAATTGACACAGCT 3338

```

1057 uHisThrLeuSerAsnThrAspAlaThrGlyLeuGlnGluIleAspA 1074
|||||
3339 GCACACTCTTCTCAGACACACAGATGACAGGCTCGAAGAAATTGACA 3388
1074 rGAlaLeuGlyTLeProGluLeuValAsnGlnGlyAlaLeuGluPro 1090
|||||
3389 GAGCTTTGGGCAATTCCTGCAACTGTCAATCAGGAGCAGCATTAAGAGCC 3438
1091 LysGlnAspAlaPheGlnGlyGlnIleAlaValMetMetAspGlnIly 1107
|||||
3439 AACACGATGCTTTCAGAGCCCAAGACAGCATGATGATGATCAGAA 3488
1107 sAlaGlyLeuTyrGlyGlnThrTyrProAlaGlnGlyProPromeGlnG 1124
|||||
3489 GGCAGATTAATATGAGACAGACATACCCAGACAGGCGCTCCAATGACAAG 3538
1124 LyGlyPheHisLeuGlnGlnGlnSerProSerPheAsnSerMetMetAsn 1140
|||||
3539 GAGGCTTTCATCTTCAGGGACATACCATCTTTAACTCTATGATGAT 3588
1141 GlnMetAsnGlnGlnGlyAsnPheProLeuGlnGlyMetHisProArgAl 1157
|||||
3589 CAGATGAACACAGCAGGCAATTTCTCTCCAGAGATGCAACCCAGAGCC 3638
1157 aAsnIleMetArgProArgThrAsnThrProLysGlnLeuArgMetGlnL 1174
|||||
3639 CAACTCATGAGACCCCGACAAACACCCCAAGCACTTAAATGACAGC 3688
1174 euGlnGlnArgLeuGlnGlnGlnGlnPheLeuAsnGlnSerArgGlnAla 1190
|||||
3689 TTCAGCAAGAGCTGCAAGGCCAGCGCTTTTGAATCAGAGCCGACAGCA 3738
1191 LeuGlnLeuLysMetGluAsnProThrAlaGlyAlaAlaValMetArg 1207
|||||
3739 CTTGAATTGAAATAAGAAACCCCTACTGCTGCTGCTGCGGTGATGAG 3788
1207 gProMetMetGlnProGlnGlnGlnGlyPheLeuAsnAlaGlnMetValAlaG 1224
|||||
3789 GCCATGATGACAGCCCGACAGAGGTTTCTTAATCCTCAAAAGCGCGCC 3838
1224 InArgSerArgGlnLeuLeuSerHisHisPheArgGlnGlnArgValAla 1240
|||||
3839 AACGACAGAGAGAGCTGCTAAGTCACTTCCGCAACAGAGGCTGGCT 3888
1241 MetMetMet.....GlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1255
|||||
3889 ATGATGATGACAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCA 3938
1255 nGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnAlaPheS 1272
|||||
3939 ACAGCAACAGCAACAGCAACAGCAGCAACAGCAACAGCAACAGCGCTTCA 3988
1272 eProProProAsnValThrAlaSerProSerMetAspGlyLeuLeuAla 1288
|||||
3989 GCCCACTCTCTAATGTGACTGCTTCCCAAGATGATGGCTTTGGCA 4038
1289 GlyProThrMetProGlnAlaProProGlnGlnPheProTyrGlnProAs 1305
|||||
4039 GGACCCCAAAATGCCCACAAAGCTCTCCGCAACAGTTCCATATCAACAAA 4088
1305 nTyrGlyMetGlyGlnGlnProAspProAlaPheGlyArgValSerSerP 1322
|||||
4089 TTATGGATGGGACAAACACAGATCCAGGCTTTGGTGAAGTGTATGTC 4138
1322 rOProAsnAlaMetMetSerSerArgMetGlyProSerGlnAsnProMet 1338
|||||
4139 CTCCCAATGCAATGATGTGCTCAAGATGGGCTCCCTCCCAAGATCCCATG 4188
1339 MetGlnHisProGlnAlaAlaSerIleTyrGlnSerSerGlnMetLysGln 1355
|||||
4189 ATGCAAACACCCGACGCTGCAATCCATCTATCAGTCTCTCAAAATGAAGG 4238

```

```

1355 yTrpProSerGlyAsnLeuAlaArgAsnSerSerPheSerGlnGlnGlnP 1372
|||||
4239 CTGGCCATCAGAGAAATTGGCCAGAGAGCTCTTTCCCGACGACAGT 4288
1372 heAlaHisGlnGlnLysnProAlaValTyrSerMetValHisMetAsnGly 1388
|||||
4289 TTGGCCACCGAGGGAAATCCTGCAGTGTATGATGTGTCACATGTAATGGC 4338
1389 SerSerGlyHisMetGlyGlnMetAsnMetAsnProMetPromeSerGln 1405
|||||
4339 AGCACTGGGCACATATGAGACAGATGAACATGAACCCCATGCCCCATGTCTCG 4388
1405 yMetProMetGlyProAspGlnLysTyrCys 1415
|||||
4389 CATGCCATAGGGTCTCTGATCAGAAATACTGTC 4419

```

seq\_name: /cgn2\_6/ptdata/1/pna/US094\_COMB.seq:US-09-440-612-3

```

seq_documentation_block:
; Sequence 3, Application US/09440612A
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; TITLE OF INVENTION: Lex M. Cowser
; FILE REFERENCE: RTS-0042
; CURRENT APPLICATION NUMBER: US/09/440,612A
; CURRENT FILING DATE: 1999-11-15
; NUMBER OF SEQ ID NOS: 49
; SEQ ID NO 3
; LENGTH: 6754
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (184)..(4422)
US-09-440-612-3

```

```

alignment_scores:
Quality: 7234.50      Length: 1427
Ratio: 5.175          Gaps: 3
Percent Similarity: 97.968      Percent Identity: 97.758

```

alignment\_block:

US-09-041-994-2 x US-09-440-612-3 ..

Align seg 1/1 to: US-09-440-612-3 from: 1 to: 6754

```

1 MetSerGlyLeuGlyGluAsnLeuAspProLeuAlaSerAspSerArgLy 17
|||||
184 ATGAGTGGATTAGAGAAAACTTGGATCCACTGGCCAGTGAATTCACGAAA 233
17 sArgLysLeuProCysAspThrProGlyGlnGlyLeuThrCysSerGlyG 34
|||||
234 ACGCAATTGGCATGTGATACACGACAGCAAGGTCTTACCTGCAATGGTG 283
34 LuLysArgArgArgGluGlnGlnSerLysTyrIleGluGluLeuAlaGlu 50
|||||
284 AAAAAGCAGAGACGGGAGCAGGAAAGTAAATATATGAGAAATGGCTGAG 333
51 LeuIleSerAlaAsnLeuSerAspIleAspAsnPheAsnValLysProAs 67
|||||
334 CTGATATCTGCCAATCTTAGTGATATATGACAAATTTCAATGTCAACCGA 383
67 pLysCysAlaIleLeuLysGluThrValArgGlnIleArgGlnIleLysG 84
|||||
384 TAAATGTGCGATTTAAAGGAACAGTAAGACAGATAGCTCAATAATAAAG 433
84 LuGlnGlyLysThrIleSerAsnAspAspAspValGlnLysAlaAspVal 100
|||||
434 AGCAAGAGAAAAAATATTTCCAAATGATGATGTTCAAAAGCCGATGTA 483
101 SerSerThrGlyGlnGlyValIleAspLysAspSerLeuGlyProLeuLe 117

```



```

|||||
484 TCTCTACAGGCGAGGAGTTATGATAAAGACTCTTAGACCGCTTTT 533
117 ULEUGINALALEUASPGLYPHELEUPHEVALVALASNAIGLUALASNI 134
|||||
534 ACTTGAGGACTTGATGGATGTTCTTATTGTGTGTAATCGAGACGGAACA 583
134 LVALPHEVALSERGLUASNVALTHRGINTYRLEUGINTYRLYSGLGLU 150
|||||
584 TTGTATTGTGATCGAAAAATGTCCACAAATACCTGCATATATAGCAAGAG 633
151 ASPLEUVALSNTHRSERVALTYRASNIIELEUHSGLUGLUASPARGLY 167
634 GACCTGGTTAAACACAGCTGTTTACATATCTTACATGAGAGACAGAAAA 683
167 SASPHELEULYSASNLEUPOLYSESTHVALASNGLYVALSERTRPT 184
|||||
684 GGATTTTCTTAAGAAATTACCAAAATCTACAGTTAATGAGTTCTCGGA 733
184 HRASNGIUPROGLINARGLINLYSERHISTHRPHASNCYSARGMETLEU 200
|||||
734 CAATGAGACCCCAAGACAAAAAGCCATTCATTATTTGCGGTATGTTG 783
201 METLYSTHRRONISASPIIELEUGLUASPIIEASNAIASERPROGLUME 217
784 ATGAAAACACACATGATATCTTGGAAGACATAAAGCCAGTCTGAAT 833
217 LARGINARGTYRGLINTHRCYSPHEALALEUSERGLNPROARGA 234
|||||
834 CGCCACAGATATGAAACAATGCAAGTCTTGCCCTCTCAGCACGACAG 883
234 LAMETMETGLUGLUGLUASPLEUGINSERCYSMETLIECYVALALA 250
|||||
884 CTATGATGAGGAGGAGGAGATTTGCAATCTGTATGATCTGTGTGCA 933
251 ARGHRTGLIETHRTHRGILYGLUARGTHRPHROSERASNPROGLUSEPH 267
934 CGCGCGATTTACTACAGGAAAGAACATTTCCATCAAAACCTGAGACTT 983
267 EILIEHTRARGHISASPLEUSERGLYLYSVALVALASNIIEASPTHASNS 284
984 TATTAACAGACATGATCTTTCAGGAAAGTTGTCATATATAGATCAAAAT 1033
284 ERLIEUARGSERSEMETARGPROGLYPHEGLUASPIIELEARGARCYSS 300
1034 CACTGAGATCTCCATGAGGCTGGCTTTGAAAGTATATATCCGAAGGTGT 1083
301 ILEGINARGPHEPHESERLIEUASNAISPGLYGINSETRPSESGINLYSAR 317
1084 ATTCAGAGATTTTCTTAAATGATGGCAGCATGCTCCCGAAACG 1133
317 GHASTYRGLINLU.....ALATYRL 324
|||||
1134 TCACATACAGAAAGTTACAGATGATGGATATTTTCCCAACAGCTTATTC 1183
324 EUSANGLYHSALAGLUTHRPROVALTYRARGPHESERLIEUALIASPGLY 340
1184 TTAAATGGCCATGCAAGAAACCCAGTATATGATTCGTGTTGGCTGATGGA 1233
341 THRILEVALTHRALAGINTHRLYSERLYSELPHEARGASNPROVALTH 357
1234 ACTATAGTACTGCACAGACAAAAAGCAAACTCTCCGAAATCTCTGTAC 1283
357 TASNASPARGHISGLYPHEVALSERTHRHISPHLEUGLARGGLUGLNA 374
1284 AAATGATGACATAGCTTGTCTCAACCCACTTCTCAGAGAAACAGA 1333
374 SNGIYTRARGPROASNPROASNPROVALGILYGLINGLYILEARGPROPRO 390
1334 ATGGATATAGCCAAACCAAAATCTGTGTGCAAGGATTTAGACCACT 1383
391 METALAGLYCYASNSESERVALGILYMETSEMETSEPROAANGI 407
|||||

1384 ATGGCTGATGACACAGTTCGGTAGCGGCATGATGTGCCAAACA 1433
407 NGLYLEUGLIMETPROSESERATGALATYRGLYLEUALIASPROSERT 424
1434 AGGCTTACAGATGCCGAGACAGAGGCGCTTAGCTTGCGAGAACCTTAGCA 1483
424 HTRHGLYGLIMETSERGLYALATRTYRGLYGLYSERSEASNIIALA 440
1484 CCACAGGCGCAGATAGTGGAGCTAGGTATGGGGTTCCAGTACATAGCT 1533
441 SERLEUTHRRPROGLYPROGLYMETGLINSERRPROSESERTYRGLINAS 457
1534 TCATTTGACCCCTGGGCCAGCATCAATCAACATCTTCCATACCAACAA 1583
457 NASRTYRGLYLEUASNETSESERPROPROHISGLYSETRPROGLYLEUA 474
1584 CAACCTATAGGCTCAACATGAGTAGCCCCCACAAGGAGAGTCGTCTTG 1633
474 LAPROASNGINGLNASNIIEMETLIESERRPROARGASNARGLYSERPRO 490
1634 CCCCACACAGCAGATATCATGATTTCTCTCTTAATCGTGGAGATCCA 1683
491 LYSILEALASERHISGLINPHESERRPROVALALAGLYVALHISSEPROME 507
1684 AAGATAGCCCTCACATCAGTTTCTCGTTGGACAGTGTGCATCTCCAT 1733
507 TALASERSEGLYASNTHRGLYASNHISSERRPHESERRSERSEUS 524
1734 GGCATCTTCTGGCAATCTGGGAACACAGCTTTTCCAGCAGCTCTCA 1783
524 ERLALEUGINALALIESERGILYVALGILYTHRSERLEULEUSERTHR 540
1784 GTGCCCTGCAAGCCATCAGTAAAGGTGGGACTTCCCTTTATCTACT 1833
541 LEUSERSERRPROGLYPROLYSLEUASPARSERRPROASNMETASNIETH 557
1834 CTGTCACTACAGGCCCCAAATTGGATATCTCTCCAAATATGATATTAC 1883
557 RGLNPROSERLYSVALSERASNGINASPSERLYSERRPROEUGLYPHER 574
1884 CCAACCAAGTAAAGTAAAGCAATCAGATTCACAAAGTCCTCGGCTTTT 1933
574 YRCYSAPRGINASNPROVALGILUSERSEMETCYSGINSERRASNFARG 590
1934 ATTCGACCAAAATCCAGTGAGAGTTCAATGTCTAGTCAAAATAGCAGA 1983
591 ASPHISLEUSERASPLYSGLUSERLYSGLYSESERVALGILYALAGI 607
1984 GATCACCTCACTGACAAAGAAAGTAAGAGACAGTGTGAGGGGCGAGA 2033
607 UASNGINARGLYPROEUGLUSERLYSGLYHSLYSLEULEUGINTL 624
2034 GAATCAAAAGGGCTCTTGGAAACAAAGTCATTAATAATTAATCAGT 2083
624 EULEUTHRCYSSERSERRASPARARGLYHISSERRSERLEUTHRASNER 640
2084 TACTTACCTGTTCTTGTATGACCGGGGCTCATTCCTCTTGACCAACTCC 2133
641 PROLEUASPSERSERCYLSYGLUSERSERRVALSERVALTHRSERPROSE 657
2134 CCCCTAATTCATTAAGTAAAGATCTTGTGTATGTTGTCACACACCCCTC 2183
657 RGLYVALSERSESERRTHRSERGLYLYVALSERSETHRSERASNMETH 674
2184 TGGAGTCTCCCTCTACATCTGAGAGAGTATCCTCAATCAATCAATATGC 2233
674 ISGLYSERLIEULEUGLUGLYHSARGTLEUHSLYSLEULEUGIN 690
2234 ATGGGTACTGTTTACAGAGAACGCCGAGATTTTGCACAAAGTTGCTCAG 2283
691 ASNGLYASNSERRPROALAGLVALALALALYSIIETHRALAGINALATHRG 707
2284 AATGGGAATTCACACAGCTGAGGTAGCCAAAGTTATCTCAGAGACCACTGG 2333

```

707 ylysaPThrSerSerIleThrSerCysGlyAspGlyAsnValValIlysg 724  
 |||||||  
 2334 GAAAGCACACACAGATATACCTTGTGGGACGGAATGTTGTCAACG 2383  
 |||||||  
 724 IuGIuGIuLeuSerProLysLysGluAsnAsnAlaLeuLeuAqTYr 740  
 |||||||  
 2384 AGGACAGCTAAgTCTTAAGAAAGAAAGAAATATATCTCTTACATAC 2433  
 |||||||  
 741 IeuleAspArgAspAspProSerAspAlaLeuSerLysGluLeuGlnr 757  
 |||||||  
 2434 CTGCTGACAGGAGATGATCTTGTGATGACACTCTTAAAGAACTACAGCC 2483  
 |||||||  
 757 oGlnValGluGlyValAspAsnLysMetSerGlnCysThrSerThrI 774  
 |||||||  
 2484 CCAAGTGAAGAGATGGATTAATAATGAGTCAGTCACACCTCCACCA 2533  
 |||||||  
 774 IeProSerSerSerGlnGluLysAspProLysIleLysThrGluThrSer 790  
 |||||||  
 2534 TTCCTAGCTCAAGTCAAGAGAAAGACCTTAAATTTAGACAGACAACT 2583  
 |||||||  
 791 GluGluGlySerGlyAspLeuAspAsnLeuAspAlaIleLeuGlyAsple 807  
 |||||||  
 2584 GAAGAGGAGATCTGAGACTTGGATTAATCTAGATGCTATTCCTGGTATCT 2633  
 |||||||  
 807 uThrSerSerAspPheTYrAsnAsnSerIleSerSerAsnGlySerHisL 824  
 |||||||  
 2634 GACTAGTCTGTGACTTTACAAATATTCATTCCTCAATGGTAGTCAATC 2683  
 |||||||  
 824 euGIyThrLysGlnGlnValPheGlnGlyThrAsnSerLeuGlyLeuLys 840  
 |||||||  
 2684 TGGGGCATAGCAACAGGTGTTCAAGAGAACTAATTCCTGGGTTTGAAA 2733  
 |||||||  
 841 SerSerGlnSerValGlnSerIleArgProTYrAsnArgAlaValSe 857  
 |||||||  
 2734 ACTTCACATCTGTGTGAGTCTATTCCTCTCATATACCGAGCACTGTC 2783  
 |||||||  
 857 rIeuAspSerProValSerValGlySerSerProProValLysAsnIleS 874  
 |||||||  
 2784 TCTGATAGCCCTGTTCTGTGGCTCAAGTCTCCAGTAAATAATATCA 2833  
 |||||||  
 874 exAlaPheProMetLeuProLysGlnProMetLeuGlyLysAsnProArg 890  
 |||||||  
 2834 GTGCTTCCCATGTTACCAAGCAACCCATGTGGGTGGGAATCCACAGA 2883  
 |||||||  
 891 MetMetAspSerGlnGluAsnTYrGlySerSerMetGlyGlyProAsnAr 907  
 |||||||  
 2884 ATGATGGATAGTCAGGAAATATGCTCAAGTATG..... 2919  
 |||||||  
 907 gasnValThrValThrGlnThrProSerSerGlyAspTrpGlyLeuProA 924  
 |||||||  
 2920 .....GGAGACTGGGCTTAACCA 2938  
 |||||||  
 924 snSerLysAlaGlyArgMetGluProMetAsnSerAsnSerMetGlyArg 940  
 |||||||  
 2939 ACTCAAGGCGCGAGAAATGGAACCTATGATTTCAAACTCCATGGGAANA 2988  
 |||||||  
 941 ProGlyGlyAspTYrAsnThrSerLeuProA-gProAlaLeuGlyGlySe 957  
 |||||||  
 2989 CCAGAGGAGATTATATCTTCTTACCAGACCTGCACCTGGGTGGCTC 3038  
 |||||||  
 957 rIleProThrLeuProLeuArgSerAsnSerIleProGlyAlaArgProV 974  
 |||||||  
 3039 TATTTCCACATTCCTCTTGGGTCTAATAGCATACAGGTGGAGCCAG 3088  
 |||||||  
 974 aIleuGlnGlnGlnGlnMetLeuGlnMetArgProGlyGluIlePro 990  
 |||||||  
 3089 TATTCACACAGACAGCAGATGCTTCAATAGAGGCTGTGTAATCCCC 3138  
 |||||||  
 991 MetGlyMetGlyAlaAsnProTYrGlyGlnAlaIleAlaSerAsnGlnLe 1007  
 |||||||  
 3139 ATGGAAATGGGGCTAATCCTATGAGCCAAAGCAGCAGCATTAACCAACT 3188

1007 uGIySerTrpProAspGlyMetLeuSerMetGluGlnValSerHisGlyT 1024  
 |||||||  
 3189 GGGTTCCTGGCCCGATGAGCATGTGTCCATGAGAACAGTTTCTCATGGCA 3238  
 |||||||  
 1024 hrcGlnAsnArgProLeuLeuArgAsnSerLeuAspAspLeuValGlyPro 1040  
 |||||||  
 3239 CTCMAATATAGCCCTCTTCTTAGGAATTCCTGGATGATCTTGTGGGCCA 3288  
 |||||||  
 1041 ProSerAsnLeuGluGlyGlnSerAspGluArgAlaLeuLeuAspGlnLe 1057  
 |||||||  
 3289 CCTTCCAACTCGGAAGAGCCAGAGTGACGAAGAGCATTTATGGACACGCT 3338  
 |||||||  
 1057 uHisThrLeuLeuSerAsnThrAspAlaThrGlyLeuGlnGluIleAspA 1074  
 |||||||  
 3339 GCACACTTCTTCAGCACACAGATGCCACAGCCCTGGAAAGAAATTGACA 3388  
 |||||||  
 1074 rGAlaLeuGlyIleProGluLeuValAsnGlnGlyGlnAlaLeuGluPro 1090  
 |||||||  
 3389 GAGCTTGGGCAATTCCTGAACTTGTCAATCAGGAGCAGGCAATTAGAGCC 3438  
 |||||||  
 1091 LysGlnAspAlaPheGlnGlnGlnGlnValAlaValMetMetAspGlnLy 1107  
 |||||||  
 3439 AAACAGAGATGCTTCCAAAGCCAAAGACAGCAGATATGATGGATCAGAA 3488  
 |||||||  
 1107 sAlaGlyLeuTYrGlyGlnThrTYrProAlaGlnGlyProProMetGlnG 1124  
 |||||||  
 3489 GCGAGATTATATGACACACATACCCACACAGGGGCTCCAAATGTCAG 3538  
 |||||||  
 1124 IyGIyPheHisLeuGlnGlnGlnSerProSerPheAsnSerMetMetAsn 1140  
 |||||||  
 3539 GAGGCTTTCATCTTCAGGACAAATCACCATCTTTTAACTCATGATGAAAT 3588  
 |||||||  
 1141 GlnMetAsnGlnGlnGlnLysAsnPheProLeuGlnGlyMetHisProArgAl 1157  
 |||||||  
 3589 CAGATGAAACACAGCAAGGCAATTTCTCTCCAAAGAAATCACCACAGAAC 3638  
 |||||||  
 1157 aAsnIleMetArgProArgTYrAsnThrProLysGlnLeuArgMetGlnL 1174  
 |||||||  
 3639 CAACATCATGAGACCCCGGACAAACACCCCAAGCACTTGAATGACACAC 3688  
 |||||||  
 1174 euGlnGlnArgLeuGlnGlnGlnGlnPheLeuAsnGlnSerArgGlnAla 1190  
 |||||||  
 3689 TTCAGCAGAGGCTGCAGGCGCAGAGTTTGTGATCAGAGCCGACAGGCA 3738  
 |||||||  
 1191 LeuGluLeuLysMetGluAsnProThrAlaGlyGlyAlaIleValMetAr 1207  
 |||||||  
 3739 CTTGAATTGAAAATGGAACCTACTGCTGTGGTGGCTGCGGTATGAG 3788  
 |||||||  
 1207 gProMetMetGlnProGlnGlnGlnPheLeuAsnAlaGlnMetValAlaG 1224  
 |||||||  
 3789 GCTTATGATGACGCCCCAGCAGGAGTTTCTTAAATGCTCAAAATGTGCCCC 3838  
 |||||||  
 1224 IArgSerArgGluLeuLeuSerHisHisPheArgGlnGlnArgValAla 1240  
 |||||||  
 3839 AACGCGACAGAGAGCGCTAATGATCATCACTTCGACAAACAGAGGGTGGCT 3888  
 |||||||  
 1241 MetMetMet.....GlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1255  
 |||||||  
 3889 ATGATGATGCACAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCA 3938  
 |||||||  
 1255 nGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1272  
 |||||||  
 3939 ACAGAACACAGCAACAGCAACAGCAGCAACAGCAACAGCAACAGCCAGCCCTTCA 3988  
 |||||||  
 1272 exProProProAsnValThrAlaSerProSerMetAspGlyLeuLeuAla 1288  
 |||||||  
 3989 GCCCACTCTCTAATGATGATGCTTCCCCAGCATGATGATGGCTTTTGGA 4038  
 |||||||  
 1289 GlyProThrMetProGlnAlaProProGlnGlnPheProTYrGlnProAs 1305  
 |||||||  
 4039 GGACCCACAATGCCACAAGCTCTCCGCAACAGTTTCCATATCAACCAAA 4088  
 |||||||  
 1305 nTYrGlyMetGlyGlnGlnProAspProAlaPheGlyArgValSerSerP 1322

```
|||||
4089 TTATGATGGACACACAGATCCAGCTTGGTGAGTGTCTACTC 4138
1322 roProAsnAlaMetSerSerArgMetGlyProSerGlnAsnProMet 1338
4139 CTCCCAATGCAATGATGCTGCAAGAAATGGCTCCCTCCAGAAATCCCATG 4188
1339 MetGlnHisProGlnAlaAlaSerIleTyrGlnSerSerGlnMetLysGln 1355
4189 ATGCACACCCCGACAGCTGCATCCATCTATCAAGTCTCAGAAATGAAGGG 4238
1355 yTrpProSerGlyAsnLeuAlaArgAsnSerSerPheSerGlnGlnLys 1372
4239 CTGGCCATCAGAGAAATTTGGCCAGGACAGCTCTTTCCAGACAGCATG 4288
1372 heaAlaHisGlnGlyAsnProAlaValTyrSerMetValHisMetAsnGly 1388
4289 TTGGCCACACAGGGAATCTCCAGTGTATAGTATGTCACATGATGATGC 4338
1389 SerSerGlyHisMetGlyGlnMetAsnMetAsnProMetProMetSerGln 1405
4339 AGCAGTGTCTACATGGGACAGATGAACATGAACCCCATGCTCATGTCTGG 4388
1405 yMetProMetGlyProAspGlnLysTyrCys 1415
4389 CATGCTATGGGTCTCTGATCAGAAATACTGC 4419
```

seq\_name: /cgn2.6/ptodata/1/pna/us096C.COMB.seq:us-09-652-123-8947

seq\_documentation\_block:

```
; Sequence 8947, Application us/09652123
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600,1186-001
; CURRENT APPLICATION NUMBER: US/09/652,123
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,135
; NUMBER OF SEQ ID NOS: 9796
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 8947
; LENGTH: 7888
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-123-8947
```

alignment\_scores:

```
Quality: 7233.50      Length: 1429
Ratio: 5.174          Gaps: 3
Percent Similarity: 97.831      Percent Identity: 97.621
```

alignment\_block:

us-09-041-994-2 x us-09-652-123-8947

Align seg 1/1 to: us-09-652-123-8947 from: 1 to: 7888

```
1 MetSerGlyLeuGlnGlyAsnLeuAspProLeuAlaSerAspSerArgLys 17
|||||
182 ATGAGTGTGATTAGAGAAACTTGATGATCCACTGGCAGTGAATTCAGGAA 231
17 SarGlyLeuProCysAspThrProGlyGlnGlyLeuThrCysSerGlyG 34
|||||
232 ACGCAAAATTCATGTGATCTCCAGGACAGAGTCTTACTGCAAGTGTG 281
34 LuLysArgArgGlnGlnGlnSerLysTyrIleGlnGlnLeuAlaGln 50
|||||
282 AAAAAGGAGACGGGACGAGAAATTAATTAATGAAGAAATGGCTAG 331
51 LeuIleSerAlaAsnLeuSerAspIleAspAsnPheAsnValLysProAs 67
|||||
```

```
332 CTGATATCTGCCAAATCTAGTATGATGACATTTCAATGTCMAACAGCA 381
67 pLysCysAlaIleLeuLysGlnThrValArgGlnIleArgGlnIleLysG 84
|||||
382 TAAATGTCCGATTTTAAAGAAACGTAAAGACAGATAGCTCAATAAAGAG 431
84 LuGlnGlySerThrIleSerAsnAspAspValGlnLysAlaLysVal 100
|||||
432 AGCAAGGAAAACTTTTCCAAATGATGATGATGTTCCAAAAAGCCGATGA 481
101 SerSerThrGlyGlnGlyValIleAspLysAspSerLeuGlyProLeu 117
|||||
482 TCTTCTACAGGGGACGAGGATTTATGATTAAGACCTCTTAGACCGCTT 531
117 uLeuGlnAlaLeuAspGlyPheLeuPheValValAsnArgGlnAlaAsn 134
532 ACTTCAGGCATTTGGATGTTCTTATTTGTTGTTGATGATGACAGCAAG 581
134 LeValPheValSerGlnAsnValThrGlnTyrLeuGlnTyrLysGlnGln 150
582 TTGATTTTGTATCAGAAATGTCAACATATCTGCATATTAATGAAGCAGAG 631
151 AsPheValAsnThrSerValTyrAsnIleLeuHisGlnGlnAspArgLys 167
632 GACCTGGTTAACACAGAGTGTATACATATCTTACATGAAGAAGACAGAA 681
167 SasPheLeuLysAsnLeuProLysSerThrValAsnGlyValSerThrPT 184
682 GGATTTTCTTAAGAAATTTACCAAAATCTACAGTTAATGAGCTTCTCTGGA 731
184 hAsnGlnProGlnArgGlnLysSerHisThrPheAsnCysArgMetLeu 200
732 CAATGAGACCCCAAGACAAAGAACCCCATACATTTAATTTGCGTATGTG 781
201 MetLysThrProHisAspIleLeuGlnLysAspIleAsnAlaSerProGln 217
782 ATGAAACACCAACATGATATCTGSAAGACATTAACGCCATCTCGAAT 831
217 TarGlnArgTyrGlnThrMetGlnCysPheAlaLeuSerGlnProArg 234
832 GCGCCAGAGATATGAAACATGACAGTGTGCTGCTCAGCCACAGCAG 881
234 IAmMetGlnGlnGlnGlyAsnLysPheGlnSerCysMetIleCysValAla 250
882 CTATGATGAGGAGAGGAGGAGATTTGCATCTTGTATATCTGTGTGCA 931
251 ArgArgIleThrThrGlyLysArgThrPheProSerAsnProGlnSerPh 267
932 CGCGCATTTACTACAGAGAAACATTTCCATCAAAACCTGAGAGCTT 981
267 eIleThrArgHisAspLeuSerGlyLysValValAsnIleAspThrAsn 284
982 TATTAACAGACATGATCTTTTCAGAAAGTGTCTCAATTAATACAAATT 1031
284 erLeuArgSerSerMetArgProGlyPheGlnAspIleIleArgArgCys 300
1032 CACTGAGATCTCCCTGAGGCTGCTGCTTGAAGATATATATCCAGAGTGT 1081
301 IleGlnArgPhePheSerLeuAsnAspGlyGlnSerThrPserGlnLysAr 317
1082 ATTCAAGACATTTTATGCTTAATATGATGGCAGTCATGCTCCAGAAAG 1131
317 gHisTyrGlnGlnLys.....AlaTyrL 324
1132 TCACATATCAAGAAAGTTACCAAGTATGATGATATTTTCCCAACAGCTAT 1181
324 euAsnGlyHisAlaGlnThrProValTyrArgPheSerLeuAlaAspGly 340
1182 TTAATGCCATGCAAGAAACCCAGTATATCAATTTCTCTTGCTGATGCA 1231
341 ThrIleValThrAlaGlnThrLysSerLysLeuPheArgAsnProValTh 357
1232 ACTATAGTACTGCACAGACAAAGCAAACTCTTCCGAAATCTGTATAC 1281
```

357 rAsnAspArgHisGlyPheValSerThrHisPheLeuGlnArgGlyGln 374  
|||||  
1282 AAATGATTCACATGGCTTGTCTCAACCCACTCTTCAGAGAGAACAGA 1331  
374 sngLyTrpArgProAsnProAsnProValGlyGlnGlyLeuArgProPro 390  
|||||  
1332 ATGGATATAGCCAAACCCAAATCTCTGGACAAAGGATTAAGACCCT 1381  
391 MetAlaGlyCysAsnSerSerValGlyGlyMetSerMetSerProAsnG 407  
|||||  
1382 ATGGCTGGATGCAACAGTTCGGTAGGGGCATGAGTATGTCGCCAAACA 1431  
407 nGlyLeuGlnMetProSerSerArgAlaTrpGlyLeuAlaAspProSer 424  
|||||  
1432 AGGCTTACAGATGCCAGACAGGGCCCTATGGCTTGCAGACCTTAGCA 1481  
424 hTrpGlyGlnMetSerGlyAlaArgTrpGlyGlySerSerHisLeuAl 440  
|||||  
1482 CCACAGGCGAGTGAAGTGAAGTATGGGGGTTCCAGTAACATAGCT 1531  
441 SerLeuThrProGlyProGlyMetGlnSerProSerSerTrpGlnAsnAs 457  
|||||  
1532 TCATTGACCCCTGGGGCAGGACATGCATCCATCTTCCACAGAACAA 1581  
457 nAsnTrpGlyLeuAsnMetSerSerProProHisGlySerProGlyLeuA 474  
|||||  
1582 CAACATAGGCTCAACATAGTAGCCCCACATGGAGTCTCGTCTTG 1631  
474 LProAsnGlnGlnAsnIleMetIleSerProArgAsnArgGlySerPro 490  
|||||  
1632 CCCCAACCAACAGCAATATCATGATTTCTCTCTGTAATCTGGAGATCCA 1681  
491 LysIleAlaSerHisGlnPheSerProValAlaGlyValHisSerProme 507  
|||||  
1682 AAGATAGCCTCAATCAGTTTCTCTCTGTTGACAGGTGCACATCTCCAT 1731  
507 tAlaSerSerGlyAsnThrGlyAsnHisSerPheSerSerSerSerLeus 524  
|||||  
1732 GGCATCTTCTGGCAATACAGGAAACACAGCTTTCACAGACTCTCTCA 1781  
524 eAlaLeuGlnAlaIleSerGlnGlyValGlyThrSerLeuLeuSerThr 540  
|||||  
1782 GTGCCCTGCAAGCCATCAGTGAAGGTGGGACTTCCCTTTATTTACT 1831  
541 LeuSerSerProGlyProLysLeuAspAsnSerProAsnMetAsnIleTh 557  
|||||  
1832 CTGTATATCACAGGCCCAATTTGATCTCTCCCAATATGATTTTAC 1881  
557 rGlnProSerLysValSerAsnGlnAspSerLysSerProLeuGlyPheT 574  
|||||  
1882 CCAAACCAGTAAAGTAAGCAATCAGGATTCAAAGAGTCTCTGGGCTTTT 1931  
574 yTrpCysAspGlnAsnProValGlnSerSerMetCysGlnSerAsnSerArg 590  
|||||  
1932 ATTTGGACCAAAATTCAGTGAAGAGTTCATATGTCAGTCAAAATACCA 1981  
591 AspHisIleLeuSerAspLysGlnSerLysGlnSerSerValGlyGlyAla 607  
|||||  
1982 GATCACCTCAGTACAAAGAAAGTAAGAGACAGTTCGAGGGGCGACA 2031  
607 nAsnGlnAlaGlyProLeuGlnSerLysGlyHisLysLysLeuLeuGln 624  
|||||  
2032 GAATCAAAAGGGTCTTTGGAAAGCAAAAGTCATAAAATTAATCTCAGT 2081  
624 euLeuThrCysSerSerAspAspArgGlyHisSerSerLeuThrAsnSer 640  
|||||  
2082 TACTTAACCTCTTCTGTGATGACCGGGGTCAATTCCTCTGACCAACTCC 2131  
641 ProLeuAspSerSerCysLysGlnSerSerValSerValThrSerProse 657  
|||||  
2132 CCCCTAGATTCAAGTTGTAAAGATCTTCTGTAGTGTACACAGCCCTTC 2181

657 rGlyValSerSerSerThrSerGlyValSerSerThrSerAsnMetH 674  
|||||  
2182 TGGAGTCTCTCTCTCTACATCTGAGAGAGTATCTTACATCCAAATATGC 2231  
674 IsGlySerLeuGlnGlnLysHisArgIleLeuHisLysLeuGln 690  
|||||  
2232 ATGGGTCACTGTACAGAGAAAGCACCGGATTTTGCACAGTTGTCGACG 2281  
691 AsnGlyAsnSerProAlaGluValAlaLysIleThrAlaGlnAlaThrG 707  
|||||  
2282 AATGGGAATTCACACACTGAGGTAGCCAAAGATTACTGCAGAAAGCCACTG 2331  
707 yLysAspThrSerSerIleThrSerCysGlyAspGlyAsnValValLysG 724  
|||||  
2332 GAAAGACACACAGCAGATTAACCTTGTGGGGCGGAAATGTTGTCAAC 2381  
724 InGlnGlnLeuSerProLysLysGluAsnAsnAlaLeuLeuArgTrp 740  
|||||  
2382 AGGAGCAGCTAAGTCTTAAGAAAGAGAGAAATATGCACTTCTTAGATAC 2431  
741 LeuLeuAspArgAspAspProSerAspAlaLeuSerLysGlyLeuGlnPr 757  
|||||  
2432 CTGCTGACAGGAGTATCTCTAGTATGCACATCTTAAGAAACTACAGCC 2481  
757 oGlnValGlyGlyValAspAsnLysMetSerGlnCysThrSerSerThrI 774  
|||||  
2482 CCAAGTGAAGAGAGTGGATATAAATGAGTCAAGTCCACACTCCACCA 2531  
774 LProSerSerSerGlnGlnLysAspProLysIleLysThrGlnThrSer 790  
|||||  
2532 TTCTTAGCTCAAGTCAAGAGAAAGCCTTAAATTAAGCAGACAGCAACT 2581  
791 GlnGlnGlySerGlyAspLeuAspAsnLeuAspAlaIleLeuGlyAspLe 807  
|||||  
2582 GAAAGAGGATCTGAGACCTTGATATCTAGATGCTATTCCTTGGTATCT 2631  
807 uThrSerSerAspPheTrpAsnAsnSerIleSerSerAsnGlySerHisL 824  
|||||  
2632 GACTAGTCTTGACTTTTACAAATTAATTCATATCCCAATTCGATGATC 2681  
824 euGlyThrLysGlnGlnValIlePheGlnGlyThrAsnSerLeuGlyLys 840  
|||||  
2682 TGGGACTAAGCAACAGGTGTTCAAGAACTAATCTCTGGGGTTGAAA 2731  
841 SerSerGlnSerValGlnSerIleArgProProTrpAsnArgAlaValSe 857  
|||||  
2732 AGTTCACACTGTGTGACGCTATTCCTCTCCATATACGAGCACTGTC 2781  
857 rLeuAspSerProValSerValGlySerSerProProValLysAsnIleS 874  
|||||  
2782 TCTGATATACCCCTGTTTCTGTGGCTCAAGTCTCCAGTAAAAAATATCA 2831  
874 eAlaIlePheProMetLeuProLysGlnProMetLeuGlyLysAsnProArg 890  
|||||  
2832 GTGCTTCCCAATGTTTACCAAAAGCAACCCATGTGGGTGGAAATCCAA 2881  
891 MetLeuAspSerGlnGluAsnTrpGlySerSerMetGlyGlyProAsnAr 907  
|||||  
2882 ATGATGATAGTACAGAAATTAATGCTCAAGTATG..... 2917  
907 gAsnValIleThrValIleGlnThrProSerSerGlyAspTrpGlyLeuPro 924  
|||||  
2918 .....GGAGACTGGGCTTAACCA 2936  
924 snSerLysAlaGlyArgMetGluProMetAsnSerAsnSerMetGlyArg 940  
|||||  
2937 ACTCAAAAGCCGCGAAGATGAACTATGAATTCAAACTCCATGGGAAGA 2986  
941 ProGlyGlyAspTrpAsnThrSerLeuProArgProAlaLeuGlyGlySe 957  
|||||  
2987 CCAGAGAGGATTAATATATCTTCTTACCCACACTGCACCTGGGTGGCTC 3036  
957 rIleProThrLeuProLeuArgSerAsnSerIleProGlyAlaArgPro 974

```

|||||
3037 TATCCACATGCGCTCTGCGTCTAATAGCATACGAGTGCAGACGAG 3086
974 aLeuGlnGlnGlnGlnGlnMetLeuGlnMetArgProGlyIlePro 990
|||||
3087 TATTGCAACAGCAGCAGCAGATGCTTCAAAATGAGCGCTGTGGAATCC 3136
991 MetGlyMetGlyAlaAsnProTyrGlyGlnAlaIleAlaSerAsnGln 1007
|||||
3137 ATGGAAATGGGGGCTAATCCCTATGCGCAAGCAGCAGCATCTAACCACT 3186
1007 uGlySerTyrProAspGlyMetLeuSerMetGlyGlnValSerHisGly 1024
3187 GGGTTCCTGGCGCCGATGCGATGTTGTCATGAGAACAGTTTCTCATGCA 3236
1024 hArgIleAsnArgProLeuLeuArgAsnSerLeuAspAspLeuValGly 1040
3237 CTCAAAATAGGCGCTCTTCTTAGGAATCCCTGATGATCTTGTGGGCGCA 3286
1041 ProSerAsnLeuGlnGlyGlnSerAspGlnArgAlaLeuLeuAspGln 1057
3287 CCTTCCAACTGAGAGCCAGAGTGCAGAAAGAGACATTTATGGACCACT 3336
1057 uHisThrLeuLeuSerAsnThrAspAlaThrGlyLeuGlnGlnIleAsp 1074
3337 GCACACTCTTCTCAGCAACACAGATGCCACAGCGCTGGAAGAAATGACA 3386
1074 rGAlaLeuGlyIleProGlnLeuValAsnGlnGlnAlaLeuGlnPro 1090
3387 GAGCTTGGGCGATCTCTGAACCTTGCATCAGGAGCGCATTTAGAGCCC 3436
1091 LysGlnAspAlaPheGlnGlnGlnAlaAlaIleMetLeuAspGln 1107
3437 AACAGGATGCTTCCAAAGGCCAAAGACAGCATTAATGATGATCAGAA 3486
1107 sAlaGlyLeuTyrGlyGlnThrTyrProAlaGlnGlyProProMetGln 1124
3487 GCGAGATTTATAGACACATATCCAGCAGCAGGGGCTCCATGACAG 3536
1124 LysGlnPheHisLeuGlnGlyGlnSerProSerPheAsnSerMetMetAsn 1140
3537 GAGGCTTTCATCTTCAGGAGCAATCACCATCTTTAACTCATATGATGAT 3586
1141 GlnMetAsnGlnGlnGlnAsnPheProLeuGlnGlnMetHisProArgAl 1157
3587 CAGATGACACAGCAGCAGCATTTTCTCTCCAGGAATGCACCCACAGAC 3636
1157 aAsnIleMetArgProArgThrAsnThrProLysGlnLeuArgMetGln 1174
3637 CAACATCATGAGACCCCGAACAAACACCCCAAGCACTTAGATGACAGC 3686
1174 euGlnGlnAlaGlyLeuGlnGlnGlnPheLeuAsnGlnSerArgGlnAla 1190
3687 TTTGAGCAGAGGCTGCAGGGCCAGAGTTTGAATCAGAGCGCAGACGCA 3736
1191 LeuGlnLeuLysMetGlnAsnProThrAlaGlyAlaIleAlaIleMetArg 1207
3737 CTGGAATTTGAAATGAAAACCTTACTGCTGCTGCTGCGGTGATGAG 3786
1207 gProMetMetGlnProGlnGlnGlnPheLeuAsnAlaGlnMetValAlaG 1224
3787 GCGATGATGACAGCCCAAGCAGGTTTCTTAATGCTCAATATGTCGCCCC 3836
1224 lnaArgSerArgGlnLeuLeuSerHisHisPheArgGlnGlnArgValAla 1240
3837 AACGACACAGAGACTCTTAAGTCATCATCTTCGACAAACAGAGGCTGCT 3886
1241 MetMetMet.....GlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1253
3887 ATGATGATGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 3936
1253 ngGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1270
|||||

```

```

3937 GCACCAACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAAA 3986
1270 lApheSerProProAsnValThrAlaSerProSerMetAspGlyLeu 1286
|||||
3987 CCTTACGCCCACTCTTAATGTGACTGCTTCCCCAGCATGATGAGGCTT 4036
1287 LeuAlaGlyProThrMetProGlnAlaProProGlnGlnPheProTyrG 1303
4037 TTGGCAGAGACCCACAAATGCCAAGCTCTCCGCAACAGTTCCATATCA 4086
1303 nProAsnTyrGlyMetGlyGlnGlnProAspProAlaPheGlyArgValS 1320
4087 ACCAAATTAATGATGGAGACAAACAGATCCAGCTTGTGCTGAGTGT 4136
1320 eRserProProAsnAlaMetMetSerSerArgMetGlyProSerGlnAsn 1336
4137 CTAGTCTCCCAATGCAATGATGTCTGCAAGAAATGGGTCCCTCCCAAT 4186
1337 ProMetMetGlnHisProGlnAlaIleAlaSerIleTyrGlnSerSerG 1353
4187 CCCATGATGCAACCCCGCAGGCTGCATCCATCTATCAGTCTCAGAAAT 4236
1353 tLysGlyTyrProSerGlyAsnLeuAlaArgAsnSerSerPheSerGln 1370
4237 GAAGGCTGGCCATCAGGAATTTGGCCAGGACAGAGCTCTTTCCACAGC 4286
1370 lnglnPheAlaHisGlnGlyAsnProAlaValTyrSerMetValHisMet 1386
4287 AGCAGTTTGGCCACAGGAGGAACTCTGACGTATGATATGTCACATG 4336
1387 AsnGlySerSerGlyHisMetGlyGlnMetAsnMetAsnProMetProme 1403
4337 AATGGCAGCAGTGTGCATGAGGAGATGAACATGAACCCCATGCCAT 4386
1403 tSerGlyMetProMetGlyProAspGlnLysTyrCys 1415
4387 GTGCGCATGGCTTATGGTCTGATCAGAAATATGTC 4423

```

seq\_name: /cgn2\_6/ptodata/1/pna/us096C\_COMB.seq:US-09-652-127-9831

seq\_documentation\_block:

```

: Sequence 9831, Application US/09652127
: GENERAL INFORMATION:
: APPLICANT: Shyjan, Andrew W.
: TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
: FILE REFERENCE: 1600.1183-001
: CURRENT APPLICATION NUMBER: US/09/652,127
: PRIOR FILING DATE: 2000-08-30
: PRIOR APPLICATION NUMBER: 60/151,134
: NUMBER OF SEQ ID NOS: 10475
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 9831
: LENGTH: 7888
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-652-127-9831

```

alignment\_scores:

```

Quality: 7233.50 Length: 1429
Ratio: 5.174 Gaps: 3
Percent Similarity: 97.831 Percent Identity: 97.621

```

alignment\_block:

US-09-041-994-2 x US-09-652-127-9831 ..

Align seq 1/1 to: US-09-652-127-9831 from: 1 to: 7888

```

1 MetSerGlyLeuGlyGlnAsnLeuAspProLeuAlaSerAspSerArgly 17
|||||
182 ATGAGTGTGATTAGGAAAACTTGTGATCCACTGGCAGTGTATTCAGCAA 231
|||||

```

17 sarGlyLeuProCysAspThrProGlyGlnGlyLeuThrCysSerGlyG 34  
 |||||  
 232 ACGCAAAATGGCATGTGATACCTCAGACAGAGTCTTACCTGCAGTGG 281  
 |||||  
 34 LuLyArgArgArgGluGlnGlnSerLysTyrlIeGluLueuLaglu 50  
 |||||  
 282 AAAACGGAGCGGAGCAGAAAGTAATATATGAAAGAAATGGCTGAG 331  
 |||||  
 51 LeuIleSerAlaSnLeuSerAspIleAspAsnPheAsnValLysProAs 67  
 |||||  
 332 CTGATATCTGCCAATCTTAGTGATATGCAATTTCAATGTCAAAACCA 381  
 |||||  
 67 pLyScysAlaIleLeuLysGluThrValArgGlnIleArgGlnIle 84  
 |||||  
 382 TAAATGTGGCATTTTAAAGAAACAGTAAGACAGATACGTAATTAAG 431  
 |||||  
 84 LuGlnGlyLysThrIleSerAsnAspAspValGlnLysAlaAspVal 100  
 |||||  
 432 ACCAAGGAAAACTATTTCCAAATGATGATGTTCCAAAAGCCGATGTA 481  
 |||||  
 101 SerSerThrGlyGlnGlyValIleAspLysAspSerLeuGlyProleu 117  
 |||||  
 482 TCTTTCTACGGGCGAGGAGTATTTGATTAAGACTCTTAGACCCGCTT 531  
 |||||  
 117 uLeuGlnAlaLeuAspGlyPheLeuPheValValAsnArgGlnAlaAsn 134  
 |||||  
 532 ACTTCAGGCAATGGATGGTTCTTATTTGGTGAATCGAGACGAAACA 581  
 |||||  
 134 leValPheValSerGluAsnValThrGlnTyrlLeuGlnTyrlLysGln 150  
 |||||  
 582 TTGTATTTGTATCAGAAATGTACACAAATACCTGCATATATAAGAA 631  
 |||||  
 151 AspLeuValAsnThrSerValTyrlAsnIleLeuHisGlnGluAspArg 167  
 |||||  
 632 GACCTGGTTAAACAAAGTCTTACATATCTTACATGAAGAAAGACGAA 681  
 |||||  
 167 saSPheLeuLysAsnLeuProLysSerThrValAsnGlyValSerTrp 184  
 |||||  
 682 GGATTTTCTTAAGAAATTTACAAAATCTACAGTTAATGAGATTCTCT 731  
 |||||  
 184 hrAsnGluProGlnArgGlnLysSerHisThrPheAsnGlySerMetLeu 200  
 |||||  
 732 CAAATGAGACCCCAAGACAAAAGCCATATTAATGGCGTATGTTGG 781  
 |||||  
 201 MetLysThrProHisAspIleLeuGlnAspIleAsnAlaSerProGlu 217  
 |||||  
 782 ATGAAACACCAATGATATTTCTGGAAGCATTAAGCCAGTCTGAAT 831  
 |||||  
 217 tarGlnArgTyrlGluThrMetGlnCysPheAlaLeuSerGlnProArg 234  
 |||||  
 832 GCGCAGAGATATGAACAAATGCAAGCTTTGCCCTCAGCCACGAG 881  
 |||||  
 234 laMetMetGlnGluGlyLysAspLeuGlnSerCysMetIleCysValAla 250  
 |||||  
 882 CTATGATGGAGGAGGGAAGATTGCAATCTGTATGATCTGTGGGA 931  
 |||||  
 251 ArgArgIleThrThrGlyLysArgThrPheProSerAsnProGluSer 267  
 |||||  
 932 CGCCCATTAACAGAGGAGAAACATTTCCATCAACCCCTGAGAGCTT 981  
 |||||  
 267 eIleThrArgHisAspLeuSerGlyLysValValAsnIleAspThrAsn 284  
 |||||  
 982 TATTACCAACATGATCTTTCAAGAAAGCTGTCAATATAGATACAAAT 1031  
 |||||  
 284 erLeuArgSerMetArgProGlyPheGluAspIleIleArgArgCys 300  
 |||||  
 1032 CACTGAGATCTTCATGAGGCGCTGCTTGAAGATATAATCCGAAGTGT 1081  
 |||||  
 301 IleGlnArgPhePheSerLeuAsnAspGlyGlnSerTrpSerGlnLys 317  
 |||||  
 1082 ATTCAAGATTTTATAGTCTAATATGATGGCAGTATGTGCCAGAAACG 1131

317 ghIstYrGlnGlu.....AlaTyrl 324  
 |||||  
 1132 TCACATATCAAGAGATTACCCAGTATGATATTTTCCCAACAGTTATTC 1181  
 |||||  
 324 euAsnGlyHisAlaGluThrProValTyrlArgPheSerLeuAlaAspGly 340  
 |||||  
 1182 TTAATGGCCATGCAGAAACCCAGTATATGATTCCTGTTGGCTGATGGA 1231  
 |||||  
 341 ThrIleValThrAlaGlnThrLysSerLysLeuPheArgAsnProValTh 357  
 |||||  
 1232 ACTATGTATGATGCACAGACAAACAAACACTTCCGAAATCTGTATAC 1281  
 |||||  
 357 rAsnAspArgHisGlyPheValSerThrHisPheLeuGlnArgGlnAla 374  
 |||||  
 1282 AAATGATGCACATGGCTTGTCTCAACCCACTTCTTCAGAGAGAAACAGA 1331  
 |||||  
 374 sngLyTyrlArgProAsnProAsnProValGlyGlnGlyLleArgPro 390  
 |||||  
 1332 ATGGATATGACCAAAACCCAAATCTGTTGGACAAAGGATTAAGACCT 1381  
 |||||  
 391 MetAlaGlyCysAsnSerSerValGlyGlyMetSerMetSerProAsnG 407  
 |||||  
 1382 ATGGCTGGATGCACAGTTCGGTAGCGGCATGATATCTCCCAACCA 1431  
 |||||  
 407 ngLyLeuGlnMetProSerSerArgAlaTyrlGlyLeuAlaAspProSer 424  
 |||||  
 1432 AGCTTACAGATGCCAGACAGAGGCTATGGCTTGGCAGACCTTAGCA 1481  
 |||||  
 424 hrThrGlyGlnMetSerGlyAlaArgTyrlGlyLysSerAsnIleAla 440  
 |||||  
 1482 CCACAGGCGAGATGAGTGAAGCTAGGTATGGGGGTTCCAGTAACATAG 1531  
 |||||  
 441 SerLeuThrProGlyProGlyMetGlnSerProSerSerTyrlGlnAsn 457  
 |||||  
 1532 TCATTTACCCCTGGGCGACAGCATGCATACCATCTCTTACCAGAACAA 1581  
 |||||  
 457 naSnTyrlGlyLeuAsnMetSerSerProHisGlySerProGlyLeuAla 474  
 |||||  
 1582 CAACATATAGCTCAACATAGTAGCCCCACATGAGGATGCTGGCTGG 1631  
 |||||  
 474 laProAsnGlnGlnAsnIleMetIleSerProArgAsnArgGlySerPro 490  
 |||||  
 1632 CCCCAACCAAGAGATATCATGATTTCTCTGTAATGTGGGAGTCCA 1681  
 |||||  
 491 LysIleAlaSerHisGlnPheSerProValAlaGlyValHisSerProMe 507  
 |||||  
 1682 AAGATATGCTCATATCAGTTTCTCTGTTCAGAGGTGCTACATCTCCAT 1731  
 |||||  
 507 laAspSerGlyAsnThrGlyAsnHisSerPheSerSerSerLeus 524  
 |||||  
 1732 GCCATCTTCTGGCAATATCTGGGAACACAGCTTTCCACACACTGTCTCA 1781  
 |||||  
 524 erAlaLeuGlnAlaIleSerGluGlyValGlyThrSerLeuSerThr 540  
 |||||  
 1782 GTGCCCTGCAGCCATCAGTGAAGGTGTGGGACTTCCCTTTATCTACT 1831  
 |||||  
 541 LeuSerSerProGlyProLysLeuAspAsnSerProAsnMetAsnIleTh 557  
 |||||  
 1832 CTGTATTCACAGGCCCAATTTGGATATCTCTCCAAATATGAATATTC 1881  
 |||||  
 557 rGlnProSerLysValSerAsnGlnAspSerLysSerProLeuGlyPheT 574  
 |||||  
 1882 CCAACCAAGTAAAGTAAGCAATCAGGATTCCAAGAGTCTCTGGGCTTTT 1931  
 |||||  
 574 yrCysAspGlnAsnProValGlySerSerMetCysGlnSerAsnSerArg 590  
 |||||  
 1932 ATTTGGACCAAAATTCAGTGAAGTTCAATGTGTACGCAAAATAGCACA 1981  
 |||||  
 591 AspHisIleuSerAspLysGluSerLysGluSerSerValGlyAlaGly 607  
 |||||  
 1982 GATCACCTCAGTGACAAAGAAAGTAAGAGACAGTGTGAGAGGGGACAGA 2031  
 |||||  
 607 uAsnGlnArgGlyProLeuGluSerLysGlyHisLysLysLeuLeuGlnL 624

```

|||||
2032 GAATCAAGGGGCTCTTGGAAAGCAAGGTCATAAAAAATTACTGCAGT 2081
624 euLeuThrCysSerSerAspAspArgLysSerSerLeuThrAsnSer 640
|||||
2082 TACTTACTCTGTTCTTCGATGACCGGGGTCATTCCTCTTGACCAACTCC 2131
641 ProLeuAspSerCysLysGlnSerSerValSerValThrSerProse 657
|||||
2132 CCCCTAGATTCAAGTTGAAAGAAATCTCTGATGTCACACGCCCTC 2181
657 rGlyValSerSerSerThrSerGlyValSerSerThrSerAsnMetH 674
|||||
2182 TGGAGCTCTCCCTCTCATCTGAGAGAGTACTCTACATCAATATGCG 2231
674 tsGlySerLeuLeuGlnGlnLysHisArgLysLeuHisLysLeuGln 690
|||||
2232 ATGGGCTCACTGTTACAAGAACGACCGGATTTGCACAAAGTTGCTCAG 2281
691 AsnGlyAsnSerProAlaGlnValAlaLysThrAlaGlnAlaThrGln 707
|||||
2282 AATGGAAATTCACCAAGCTGAGGTAGCCAAAGATTACTGCAAGGCCACTGG 2331
707 LysAspThrSerSerThrSerCysGlyAspGlyAsnValAlaLysG 724
|||||
2332 GAAAGACACACAGAGTAACTTCTTGGGGAGCGAAATGTTGCAAGC 2381
724 lngLglnLeuSerProLysLysLysGlnAsnAlaLeuLeuArgTyr 740
|||||
2382 AGGAGCGAGTAAAGTCTTAAGAAAGAGAAATTAATGCACTTCTTAAGTAC 2431
741 LeuLeuAspThrGAspAspProSerAspAlaLeuSerLysGlnLeuGlnPr 757
|||||
2432 CTGCTGACAGGAGATGATCCTAGTATGACACTCTCTAAAGACTACAGCC 2481
757 OGlnValGlnGlyValAspAsnLysMetSerGlnCysThrSerSerThrI 774
|||||
2482 CCAAGTGGAAAGAGTGGATATAATAATGAGTCACTGCACGCCACTGCCACA 2531
774 leProSerSerSerGlnGlnLysAspProLysIleLysThrGlnThrSer 790
|||||
2532 TTCTACTCTCAAGTCAAGAGAAAGACCTTAATAATTAAGACAGACAGCAAGT 2581
791 GlnGlnGlySerGlyAspLeuAspAsnLeuAspAlaIleLeuGlnLysPle 807
|||||
2582 GAAGAGGGATCTGGAGACTTGGATTAATCTAGATCTATTTCTGGTGATCT 2631
807 urThrSerSerAspPheTyrAsnAsnSerIleSerSerAsnGlySerHisL 824
|||||
2632 GACTAGTTCTGACTTTACATTAATTCATATCCTCAAAATGTTGATCATC 2681
824 euGlyThrLysGlnGlnValPheGlnGlnLysThrAsnSerLeuGlyLeuLys 840
|||||
2682 TGGGAGACTAAGCAACAGGTGTTTCAAGAACTAATTTCTTGCGTTTGAAA 2731
841 SerSerGlnSerValGlnSerIleArgProProLysArgAlaValSer 857
|||||
2732 AGTTCAACAGTCTGTCAGTCTATTGCTCTCCATATAACGAGACAGTCTC 2781
857 rLeuAspSerProValSerValGlySerSerProProValLysAsnLys 874
|||||
2782 TCTGGATAGCCCTGTTCTGTTGGCTCAAGTCCCAAGTAAAAATATNCA 2831
874 eAlaAlaPheProMetLeuProLysGlnProMetLeuGlyGlyAsnProArg 890
|||||
2832 GTGCTTTTCCCATGTATCAAAAGCAACCCATGTGGGAGGAATTCAGAA 2881
891 MetMetAspSerGlnGlnLysAsnTyrGlySerSerMetGlyGlyProAsnAr 907
|||||
2882 ATGATGATGATGTCAGAAAATTAATGCGCTCAAGTATG..... 2917
907 GAsnValThrValThrGlnThrProSerSerGlyAspThrGlyLeuProA 924
|||||
|||||
2918 ..... GGAGACTGGGGCTTACCAA 2936
924 snSerLysAlaGlyArgMetGluProMetAsnSerAsnSerMetGlyArg 940
|||||
2937 ACTCAAGGCCGCCAGATGAGAACCTATGAATTAACCTCATGGGAGAGA 2986
941 ProGlyGlyAspTyrAsnThrSerLeuProArgProAlaLeuGlnGlyLys 957
|||||
2987 CCAGGAGGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3036
957 rIleProThrLeuProLeuArgSerAsnSerIleProGlyAlaArgProV 974
|||||
3037 TATTCACCATTAATGCTCTGTCGCTTAATAGCATTCAGAGTCCGAGACAG 3086
974 aLleuGlnGlnGlnGlnGlnMetLeuGlnMetArgProGlyGlnIlePro 990
|||||
3087 TATTGCAACACACAGACAGAGATGCTTCAAAAGAGCCCTGGGGAATATCCC 3136
991 MetGlyMetGlyAlaAsnProTyrGlyGlnAlaAlaLysSerAsnGlnI 1007
|||||
3137 ATGGGAATGGGGCTTAATCCTTAATGCGCACAGACATCTTAACCAACT 3186
1007 uGlySerThrProAspGlyMetLeuSerMetGlnGlnValSerHisGlyT 1024
|||||
3187 GGGTTCTCGGCCGATGGCATGTTGTCATGGAACAAGTTTCTCATGSCA 3236
1024 hrGlnAsnArgProLeuLeuArgAsnSerLeuAspAspLeuValGlyPro 1040
|||||
3237 CTCAAAATAGGCTCTTCTTAAGAAATCCCTGGAATGATCTTGTGGGCCA 3286
1041 ProSerAsnLeuGlnGlyGlnSerAspGlnArgAlaLeuLeuAspGlnI 1057
|||||
3287 CCTTCCAACTCGAAGCCAGAGTGCAGAAAGACATTAATGACCAAGCT 3336
1057 uHisThrLeuLeuSerAsnThrAspAlaThrGlyLeuGlnGlnIleAsp 1074
|||||
3337 GCACACTCTTCTCGAACACACAGATGCCACAGGCCCTGGAAAGAAATTGCA 3386
1074 rGAlaLeuGlyIleProGlnLeuValAsnGlnGlnGlnAlaLeuGlnPro 1090
|||||
3387 GAGCTTTGGGCACTTCGAACTTGTCAATCAAGGACAGGCTTGTAGAGCC 3436
1091 LysGlnAspAlaPheGlnGlnGlnGlnAlaAlaValMetMetAspGlnI 1107
|||||
3437 AAACAGGATGCTTTCCAAAGCCCAAGAACACAGCATTAATGATGATCACA 3486
1107 saLagLysLeuTyrGlnGlnThrTyrProAlaGlnGlnLysProProMetGln 1124
|||||
3487 GGCAGGATTAATATGACAGACATACCCAGACAGGGGCTCCAAATGCAAG 3536
1124 LysGlyPheHisLeuGlnGlnGlnSerProSerPheAsnSerMetMetAsn 1140
|||||
3537 GAGGCTTTCACTTCAGGAGCAATCACCATCTTTTAATCTCATATGATGAT 3586
1141 GlnMetAsnGlnGlnGlnLysAsnPheProLeuGlnGlnGlnMetHisP 1157
|||||
3587 CAGATGAACCAAGCAAGCAATTTTCTCTCCAAAGAAATGCACCCAGAGC 3636
1157 aaSnIleMetArgProArgThrAsnThrProLysGlnLeuArgMetGlnI 1174
|||||
3637 CAACATATATGAGACCCGGAGCAAAACCCCAACCACTTAATGAATGACGC 3686
1174 euGlnGlnArgLeuGlnGlnGlnGlnPheLeuAsnGlnSerArgGlnAla 1190
|||||
3687 TTCAGCAGAGGCTCAGAGGCCAGAGATTTTGAATTCAGAGCCGACAGCA 3736
1191 LeuGlnLeuLysMetGlnAsnProThrAlaGlyLysAlaAlaValMet 1207
|||||
3737 CTTGAATTTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 3786
1207 gProMetMetGlnProGlnGlnGlnGlnPheLeuAsnAlaGlnMetValaG 1224
|||||
3787 GCTTATGATGACGCCCAAGAGGTTTCTTAATGCTCAATATGCTCGGCC 3836

```

```
1224 InaArgSerArgGluLeuLeuSerHisHisPheArgGlnGlnInaArgValAla 1240
|||||
3837 AACGACAGACGAGAGCTGCTAACTCACTTCGACAAACAGAGGGGTGCT 3886
1241 MetMetMet.....GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1253
|||||
3887 ATGATGATGTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 3936
1253 ngInGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1270
|||||
3937 GCAGCAACAGCAACAGCAACAGCAACAGCAACAGCAACAGCAACAGCAAC 3986
1270 IaPheSerProProProProValIthrIaSerProSerMetAspGlyLeu 1286
|||||
3987 CCTCAGCCACCTCCTAATGTGACTGCTTCCCGCAGCATGGATGGGCTT 4036
1287 LeuAlaGlyProThrMetProGlnAlaProProGlnGlnPheProTyrGln 1303
|||||
4037 TTGGGAGGACCCACAATGCCACAAGCTCCTCCGCAACAGTTTCCATATCA 4086
1303 nProAsnTyrGlyMetGlyGlnGlnIProAspProAlaPheGlyArgValS 1320
|||||
4087 ACCAATTTATGSAATGGGACCAACACAGATCCAGCTTTGGTCCAGTGT 4136
1320 eArgSerProProAsnAlaMetMetSerSerArgMetGlyProSerGlnAsn 1336
|||||
4137 CTAGTCCTCCCAATGATGATGCTGCTCAAGATGGGTCCCTCCCAAGAT 4186
1337 ProMetMetGlnHisProGlnAlaIaSerIleTyrGlnSerSerGlnMet 1353
|||||
4187 CCCATGATGCAACACCCGAGGCTGATCATCTATCATAGTCTCCAGAAAT 4236
1353 tLysGlyTyrProSerGlyAsnLeuAlaArgAsnSerSerPheSerGlnG 1370
|||||
4237 GAAGGGCTGGCCATCAGGAATTTGGCCGGAACAGCTCTTTTCCACAC 4286
1370 IuGlnPheAlaHisGlnGlyAsnProAlaValTyrSerMetValHisMet 1386
|||||
4287 AGCAGTTTGGCCACACAGGGGAATCCTGAGTATAGTATGCTGCACAG 4336
1387 AsnGlySerSerGlyHisMetGlyGlnMetLysMetAsnProMetProMet 1403
|||||
4337 AATGGCAGCAGTGTGTCATGAGGACAGATGAAACATGAAACCCCATGGCC 4386
1403 tSerGlyMetProMetGlyProAspGlnLysTyrCys 1415
|||||
4387 GTCGTGGCATGGCTATGGGTCTGATCAGAAATATCTGC 4423
seq_name: /cgrn2_6/ptcodata/1/pna/us09-09-652-355-11058
seq_documentation_block:
: Sequence 11058, Application US/09652355
: GENERAL INFORMATION:
: APPLICANT: Shyjan, Andrew W.
: TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
: TITLE OF INVENTION: THEREFOR
: FILE REFERENCE: 1600.1192-001
: CURRENT APPLICATION NUMBER: US/09/652.355
: PRIOR FILING DATE: 2000-08-30
: PRIOR APPLICATION NUMBER: 60/151.136
: NUMBER OF SEQ ID NOS: 11227
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 11058
: LENGTH: 7888
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-652-355-11058
alignment_scores:
Quality: 7233.50 Length: 1429
```

```
Ratio: 5.174 Gaps: 3
Percent Similarity: 97.831 Percent Identity: 97.621
alignment_block:
US-09-041-994-2 x US-09-652-355-11058 ..
Align seg 1/1 to: US-09-652-355-11058 from: 1 to: 7888
1 MetSerGlyLeuGlyGluAsnLeuAspProLeuAlaSerAspArgGly 17
|||||
182 ATGAGTGGATTTAGGAAACCTTGATCCATCGCCAGATTCAGAA 231
17 sArgLysLeuProCysAspThrProGlyGlnGlyLeuThrCysSerGlyG 34
|||||
232 ACGCAAAATGGCATGTGATCTCCAGACAAAGTCTTACTCTCAGATGGG 281
34 LuLysArgArgArgGluGlnGlnLysSerTyrIleGluGluLeuAlaGlu 50
|||||
282 AAAAAGGAGAGCGGAGCAGAAAGTAATATATTTGAAGAAATGGCTGAG 331
51 LeuIleSerAlaAsnLeuSerAspIleAspAsnPheAsnValLysProAs 67
|||||
332 CTGATATCTGCCAATCTTGATATGTGACATTTCAATGTCAAACAGCA 381
67 pLysCysAlaIleLeuLysGluThrValArgGlnIleArgGlnIleLysG 84
|||||
382 TAAATGTGGCATTTTAAAGAAACAGTAAGACAGATACGCTCAAAATAAAG 431
84 LuGlnGlyLysThrIleSerAsnAspAspAspValGlnLysAlaAspVal 100
|||||
432 ACCAAGGAAACCTATTTCCATGATGATGATGTTCCAAAGCCCATGTGA 481
101 SerSerThrGlyGlnGlyValIleAspLysAspSerLeuGlyProLeuLe 117
|||||
482 TCTTGTACAGGCGCAGGAGTATGTGATAAGACTCCTTAGACCCGCTTT 531
117 tLeuGlnAlaLeuAspGlyPheLeuPheValValAsnArgGluAlaAsnI 134
|||||
532 ACTTCAGGCAATGGATGGTTCCTATTTGCTGTGATCAGACGGAACAA 581
134 lValAlaPheValSerGluAsnValIThrGlnTyrLeuGlnIleTyrLysGln 150
|||||
582 TTGTATTTTATCAGAAATGTCACACAAATACCTGCAATATTAAGCAAG 631
151 AspLeuValAsnThrSerValTyrAsnIleLeuHisGluGluAspArgLys 167
|||||
632 GACCTGGTTAAACAAGTGTATACAAATATCTTACATGAAGAGACAGAA 681
167 sAspPheLeuLysAsnLeuProLysSerThrValAsnGlyValSerTyrP 184
|||||
682 GGAATTTTCTTAAGAAATTTACAAAATCTACAGTTAATGAGATTTCTCG 731
184 hPAsnGluProGlnArgGlnLysSerHisThrPheAsnCysArgMetLeu 200
|||||
732 CAATGAGACCCCAAGACAAAAGCAATTAATTAATGCGCGTATGTTG 781
201 MetLysThrProHisAspIleLeuGluAspIleAsnAlaSerProGluMe 217
|||||
782 ATGAAAACACACATGATGATCTGGAAGACATAAAGCCAGAGCTGAAT 831
217 tArgGlnArgGlyGluThrMetGlnCysPheAlaLeuSerGlnProAlaG 234
|||||
832 GCGCCAGAGATATGAACAATGACAGTCTTCCCTGCTTCAGCCAGCG 881
234 lMetMetGluGluGlyGluAspLeuGlnSerCysMetIleCysValAla 250
|||||
882 CTATGATGAGGAGAGGGAAGATTTGCAATCTTGATGATCTGTGTGCA 931
251 ArgArgIleThrThrGlyLysArgThrPheProSerAsnProGluLysArg 267
|||||
932 CCGCGCATTAACAAGAGAAAGAACATTTCCATCAAAACCTGAGAGCTT 981
267 eIleThrArgHisAspLeuSerGlyLysValValAsnIleAspThrAsnS 284
```



```

|||||
982 TATTACACACATGATCTTTACAGAAAGGTGTGCATATATAGATACAAATT 1031
284 eTLeuArgSerSerMetArgProGlyPheGluAspIleIleArgArgCys 300
1032 CACTGATGATCTTCATGAGGCGCTGTGGAAGATATTAATCCGAAAGGTGT 1081
301 ILeuIleArgPhePheSerLeuAsnAspGlyGlnSerTyrSerGlnIleAsn 317
1082 ATTACAGAGATTTTATAGTCTAAATGATGGGACATGTGTCCTCCAGAAACG 1131
317 gHISyrGlnGlu.....AlaTyrL 324
1132 TCACATATCAAGAAGTTACCAAGTATGATATTTCCCAACAGCTTATC 1181
324 euAsnGlyIleIleAspGluThrProValTyrArgPheSerLeuAlaAspGly 340
1182 TTAAATGGCCATGCAGAAACCCAGATATCGATTCCTGTGGCTGATGCA 1231
341 ThrIleValThrAlaGlnThrLysSerLysLeuPheArgAsnProValThr 357
1232 ACTATATGATCTGCACAGACAAACAAACAACTTCCGAAATCCTGTATC 1281
357 rAsnAspArgHISglyPheValSerThrHISpheLeuGlnArgGluGlnA 374
1282 AATATGATCGACATGGCTTGTCTCAACCCACTTCCTTCAGAGAGAACAGA 1331
374 sngLYTyrArgProAsnProAsnProValGlyGlnGlyIleArgProPro 390
1332 ATGGATATGACCAAAACCAAAATCCTGTGTGGACAAAGGATTTAGACCACT 1381
391 MetAlaGlyCysAsnSerSerValGlyLysMetSerMetSerProAsnG 407
1382 ATGGCTGATGCAACAGTTGCGTAGGCGCATGATGATGTCCGCAAAACCA 1431
407 nGlyLeuGlnMetProSerSerArgAlaTyrGlyLeuAlaAspProSerT 424
1432 AGGCTTACAGATGCCGACGACGAGCGCCATGCGCTTGGCAGACCCCTAGA 1481
424 hTyrGlyGlnMetSerGlyAlaArgTyrGlyLysSerSerAsnIleAla 440
1482 CCACAGGCGACGATGAGTGGAGTAGTATGGGCGTCCAGTAAACATAGCT 1531
441 SerLeuThrProGlyProGlyMetGlnSerProSerSerTyrGlnAsnS 457
1532 TCATTGACCCCTGGGCGACGATGCAATCACCATTCTTACACAGAACAA 1581
457 nAsnTyrGlyLeuAsnMetSerSerProProHISGlySerProGlyLeuA 474
1582 CAAGCTATAGGCTCAACATGAGTAGCCGCCACATGGAGTCCCTGGTCTTG 1631
474 lApProAsnGlnGlnAsnIleMetIleSerProArgAsnArgGlySerPro 490
1632 CCCCACAAACAGCAGAAATATCATGATTTCTCTCGTAAATCGTGGAGATCCA 1681
491 LysIleAlaSerHISGlnPheSerProValAlaGlyValHISerProMe 507
1682 AAGATAGCCTCAATCAGTTTCTCTGTGGAGGTGCACACTCTCCAT 1731
507 tAlaSerSerGlyAsnThrGlyAsnHISerPheSerSerSerSerLeuS 524
1732 GGCATCTTCTGGCAATACCTGGAGACACAGCTTTCACAGACGCTCTCTA 1781
524 eTAlaLeuGlnAlaIleSerGlyGlyValGlyThrSerLeuLeuSerThr 540
1782 GTGCCCTGCAGACCATCACTGAAGGTGTGGGACCTCCCTTTATCTACT 1831
541 LeuSerSerProGlyProLysLeuAspAsnSerProAsnMetAsnIleH 557
1832 CTGTCTATCCAGGCCCAAAATTTGATTAATCTCCCAATATGAAATATAC 1881
557 rGlnProSerLysValSerAsnGlnAspSerLysSerProLeuGlyPheT 574
|||||

```

```

1882 CCAACCAAGTAAAGTAAAGCAATCAGATTCCAAGAGTCCCTGGGCTTTT 1931
574 yTCysAspGlnAsnProValAlGlySerSerMetCysGlnSerAsnSerArg 590
1932 ATTCGACCAAAAATCCAGTGGAGAGTTCAATGTGTAGTCAAAATATACAGA 1981
591 AspHISLeuSerAspLysGlnSerLysGlnSerSerValGlyValAlaG 607
1982 GATCACCTCAGTGACAAAGAAAGTAAGAGACAGCTTGAGAGGGGCACA 2031
607 uAsnGlnArgGlyProLeuGlnSerLysGlyHISLysLysLeuLeuGlnL 624
2032 GAATCAAGGGGCTCTTGGAAAGCAAGGTCAATAAAAAATTACTGCAGT 2081
624 euLeuThrCysSerSerAspArgArgGlyHISerSerLeuThrAsnSer 640
2082 TACTTACCTGTTCTTGATGACCCGGGTCATTCCTCTGACCAACTCC 2131
641 ProLeuAspSerSerCysLysGlnSerSerValSerValThrSerProSe 657
2132 CCCCTGATTCAGATTGTAAAGATCTTCTGTAGTGTACCAACGCCCTTC 2181
657 rGlyValSerSerSerThrSerGlyValAlaSerSerThrSerAsnMetH 674
2182 TGGAGTCTCCTCCTTACATCTGGAGAGATATCCTTACATCAATATATGC 2231
674 iSGLysrLeuLeuGlnGlyLysHISArgIleLeuHISLysLeuLeuGln 690
2232 ATGGGTCACCTGTTTACAAAGAAAGCACCGGATTTTGCACAAAGTTGCGCAG 2281
691 AsnGlyAsnSerProAlaGlnValAlaLysIleThrAlaGlnAlaThrG 707
2282 AATGGAAATTCACCACTGATGACCAAGATTTACTGCAAGAACCACTGG 2331
707 yLysAspThrSerSerIleThrSerCysGlyAspGlyAsnValValLysG 724
2332 GAAAGACACACAGACATATAACTTCTGTGGGACGGAATGTTGTCAACG 2381
724 lngLglnLeuSerProLysLysLysGlnAsnAsnAlaLeuLeuArgTyr 740
2382 AGGAGCAGCTAAGTCTTAAGAAAGAGAAATATATCATCTTGTAGATAC 2431
741 LeuLeuAspArgAspAspProSerAspAlaLeuSerLysGlnLeuGlnPr 757
2432 CTGCTGACAGAGGATATCTAGTATGCACTCTTAAAGAACTACAGGC 2481
757 oGlnValGlyGlyValAspAsnLysMetSerGlnCysThrSerSerThrI 774
2482 CCAAGTGAAGAGGTGGATTAATAATGATGATGATGATGATGATGATGATG 2531
774 lApProSerSerSerGlnGlyLysAspProLysIleLysThrGlyThrSer 790
2532 TTCTCTGATCAAGTCAAGTCAAGAAAGACCTTAAATTTAAAGACAGACAACT 2581
791 GluGlnGlySerGlyAspLeuAspAsnLeuAspAlaIleLeuGlySple 807
2582 GAAGAGGATCTGGAGACTGTGGATTAATCTTAAGTATCTTGGTATAT 2631
807 uThrSerSerAspPheTyrAsnAsnSerIleSerSerAsnGlySerHISL 824
2632 GACTATGTTCTGACTTTTCAATATAATTCATATCCTCAATATGATGATG 2681
824 euGlyThrLysGlnGlnValPheGlnGlyThrAsnSerLeuGlyLeuLys 840
2682 TGGGACACTAAGCAACAGGTGTTCAGAGAAACATTAATCTCTGGGTTGAAA 2731
841 SerSerGlnSerValGlnSerIleArgProProLysArgArgAlaValSe 857
2732 AGTTCACTCTGTGACGTATATGCTCTCCATATATACCGAGCAGGTGTC 2781
857 rLeuAspSerProValSerValGlySerSerProProValLysAsnIleS 874
2782 TCTGGATACCCCTGTTTCTGTGGTCAAGTCTCTCAGTAAATAATATCA 2831

```

874 eraIaPheProMetLeuProLySGlnProMetLeuGlyGlyAsnProArg 890  
 |||||  
 2832 GTGCTTCCCATGTACCAAAAGCAACCATGTTGGGTGGGAATCCAGA 2881  
 891 MetMetAspSerGlnIuAsnProThrAlaGlyGlySerMetGlyProAsnAr 907  
 |||||  
 2882 ATGATGGATGTCTAGAAATATATGGCTCAAGTATG..... 2917  
 907 GAsnValThrValThrGlnThrProSerSerGlyAspTrpGlyLeuProA 924  
 |||||  
 2918 .....GAGACTGGGGCTTACCAG 2936  
 924 snSerLysAlaGlyArgMetGluProMetLAsnSerAsnSerMetGlyArg 940  
 |||||  
 2937 ACTCAAAAGCGCGCAGATGGAACCTATGAATCAAACTCATCGGAGAGA 2986  
 941 ProGlyLysAspTrpAsnThrSerLeuProArgProAlaLeuGlyGlySe 957  
 |||||  
 2987 CCAAGAGAGATTAATTAATCTTACCCAGACTGCACCTGGGTGGCTC 3036  
 957 rIleProThrLeuProLeuArgSerAsnSerIleProGlyAlaArgProV 974  
 |||||  
 3037 TATTCACACATGGCTCTTGGCTTAATAGCATACAGGTGCGAGACAG 3086  
 974 aLeuGlnGlnGlnGlnMetLeuGlnMetArgProGlyGlyIlePro 990  
 |||||  
 3087 TATTGCAACAGCAGCAGCAGATGCTCAAAATGAGCGCTGTGGAATCCC 3136  
 991 MetGlyMetGlyAlaAsnProTrpGlyGlnAlaAlaAlaSerAsnGlnIle 1007  
 |||||  
 3137 ATGGGAATGGGGGCTAATCCCTATGGCCAGACAGCAGCATTTAAACCACT 3186  
 1007 uGlySerTrpProAspGlyMetLeuSerMetGluGlnValSerHisGlyT 1024  
 |||||  
 3187 GGGTCTCTGGCCGATGCAATGTTGTCATGGAACAAGTTTCTCATGGCA 3236  
 1024 hrgIAsnAsnArgProLeuLeuArgAsnSerLeuAspAspLeuValGlyPro 1040  
 |||||  
 3237 CTCAAAATAGGCTCTTCTTGAAGAAATCCCGATGATCTTGTGGGCCA 3286  
 1041 ProSerAsnLeuGlnGlyGlnSerAspGlyuArgAlaLeuLeuAspGlnIle 1057  
 |||||  
 3287 CCTTCCCAACCTCGAAGGCCAGAGCAGCAAGACATTTATGGACCAAGCT 3336  
 1057 uHisThrLeuLeuSerAsnThrAspAlaThrGlyLeuGlnGluIleAspA 1074  
 |||||  
 3337 GCACACTCTTCTCAGCAACACAGATGCCACAGCGCTGGAAGAAATGACA 3386  
 1074 rglAlaLeuGlyIleProGluLeuValAsnGlnGlyGlnAlaLeuGluPro 1090  
 |||||  
 3387 GAGCTTTGGGCAATTCGTGAACCTTGTCAATCAGGGACAGGCATTAGACCCC 3436  
 1091 LysGlnAspAlaPheGlnGlyGlnGlnAlaAlaValMetMetAspGlnIle 1107  
 |||||  
 3437 AAACAGCATGCTTCCAAAGGCCAAGAACAGCAGCATTAATGATGATCGA 3486  
 1107 saIaGlyLeuTrpGlyGlnThrTrpProAlaGlnGlyProProMetGlnG 1124  
 |||||  
 3487 GCGAGGATTTATGTGACAGACATACCCAGCAGGGGCTTCATATGCAAG 3536  
 1124 LysGlyPheHisLeuGlnGlyGlnSerProSerPheAsnSerMetMetLAsn 1140  
 |||||  
 3537 GAGGCTTTCACTTCAGGAGCAATCACCATCTTTAACTCATATGATGANT 3586  
 1141 GlnMetAsnGlnGlnGlnLysAsnPheProLeuGlnGlnLysMetHisProArgAl 1157  
 |||||  
 3587 CAGATGAACCGCAGAGCAATTTCTCTCCAAAGAAATGCAACCCACAGAGC 3636  
 1157 aAsnIleMetArgProArgThrAsnThrProLySGlnLeuArgMetGlnL 1174  
 |||||  
 3637 CAACATCATGAGACCCGAGCAAAACACCCCAACCACTTAGAATGCAAGC 3686

1174 euGlnGlnAthrGleuGlnGlnGlnPheLeuAsnGlnInserArgGlnAla 1190  
 |||||  
 3687 TTCAGCAGAGGCTGCGAGGGCCAGCATGTTTGAATGAGACCCGACAGCA 3736  
 1191 LeuGluLeuLysMetGluAsnProThrAlaGlyGlyAlaAlaValMetAr 1207  
 |||||  
 3737 CTTCGAATTTGAATGAAATGAAACCCCTACTGCTGTGTGTCTCTCGGGTATGAG 3786  
 1207 gProMetMetGlnProGlnGlnGlnPheLeuAsnAlaGlnMetValAlaG 1224  
 |||||  
 3787 GCCTATGATGAGCCCGCAGCAGGGGTTTCTTAATGCTCAATAGTGTGCCCC 3836  
 1224 LnaArgSerArgGluLeuLeuSerHisHisPheArgGlnGlnArgValAla 1240  
 |||||  
 3837 AACGACAGAGAGAGCTCTTAATGTCATCATCTCCGACACAGAGGGTGGCT 3886  
 1241 MetMetMet.....GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1253  
 |||||  
 3887 ATGATGATGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 3936  
 1253 nGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1270  
 |||||  
 3937 GCAGCAACAGCAACAGCAACAGCAACAGCAACAGCAACAGCAACAGCA 3986  
 1270 LAsPheSerProProProAsnValThrAlaSerProSerMetAspGlyLeu 1286  
 |||||  
 3987 CTTTCACGCCCACTCTTAATGTGACTCTTCCCGCAGCATGAGTGGGCTT 4036  
 1287 LeuAlaGlyProThrMetProGlnAlaProProGlnGlnPheProTrpGly 1303  
 |||||  
 4037 TTGGCAGAGCCCAACATGCCAAGCTCTCTCCGCAACAGTTTCCATATCA 4086  
 1303 nProAsnTrpGlyMetGlyGlnGlnProAspProAlaPheGlyArgValS 1320  
 |||||  
 4087 ACCAAATTTATGGAATGGGACAAACAGATCCAGCTTGTGTGAGTGT 4136  
 1320 eSerProProAsnAlaMetMetSerSerArgMetGlyProSerGlnAsn 1336  
 |||||  
 4137 CTAGTCTCTCCCAATGCAATGATGTGTGTCAGAAATGGTCTCCCAAAAT 4186  
 1337 ProMetMetGlnHisProGlnAlaAlaSerIleTrpGlnSerSerGlnIle 1353  
 |||||  
 4187 CCCATGATGACACACCGCAGGCTGCATTCATCATAGTACGTCTAGAAAT 4236  
 1353 tLysGlyTrpProSerGlyAsnLeuAlaArgAsnSerSerPheSerGln 1370  
 |||||  
 4237 GAAGGCTGGCCATCAGAAATTTGGCCAGGAAACAGCTCTTTCCCAAGC 4286  
 1370 lngInPheAlaHisGlnGlnLysAsnProAlaValLysSerMetValHisMet 1386  
 |||||  
 4287 AGCAGTTTGGCCACAGGGGGAATCTTCAGATGTAATGATGTGTGCACATG 4336  
 1387 AsnGlySerSerGlyHisMetGlyGlnMetLAsnMetAsnProMetProMe 1403  
 |||||  
 4337 AATGCGACGATGCTCATATGAGACATGAAATGAACCCCAATGCCCAT 4386  
 1403 tSerGlyMetProMetGlyProAspGlnLysTrpCys 1415  
 |||||  
 4387 GTCTGCATGCGCTATGGGTCTGATCAGAAATACTGC 4423  
 seq\_name: /cgn2\_6/plodata/1/pna/us096D\_COMB.seq:us-09-663-693-1028  
 seq\_documentation\_block:  
 ; Sequence 1028, Application us/09663693  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Geating, David P.  
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES  
 ; FILE REFERENCE: 1600.1195-001  
 ; CURRENT APPLICATION NUMBER: US/09/663,693  
 ; CURRENT FILING DATE: 2000-09-21  
 ; PRIOR APPLICATION NUMBER: 60/154,986  
 ; PRIOR FILING DATE: 1999-09-21

```

; NUMBER OF SEQ ID NOS: 1340
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 1028
; LENGTH: 7888
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-663-693-1028

```

```

alignment_scores:
    Quality: 7233.50      Length: 1429
    Ratio: 5.174          Gaps: 3
Percent Similarity: 97.831 Percent Identity: 97.621

```

```

alignment_block:
US-09-041-994-2 x US-09-663-693-1028

```

```

Align seg 1/1 to: US-09-663-693-1028 from: 1 to: 7888

```

```

1 MetSerGlyLeuGlyLysAsnLeuAspProLeuAlaSerAspSerArgly 17
   |||||
182 ATGAGTGGATTAGAGAAAACCTGGATCCACTGGCCAGTATTCACGAAA 231
   |||||
17 sarGlyLeuProCysAspThrProGlyGlnGlyLeuThrCysSerGly 34
   |||||
232 ACCCAATTGCCATGTGATCTCAGACACAAGCTTACCTGCGAGTGTG 281
   |||||
34 LulSarGArGArGluGlnGlnSerLysTyrIleGlnGluLeuAlaGlu 50
   |||||
282 AAAAACGGAACGGAGACGAAAGTAATATATGAAAGATTGGCTGAG 331
   |||||
51 LeuIleSerAlaAsnLeuSerAspIleAspAsnPheAsnValLysProAs 67
   |||||
332 CTGATATCTGCCAATCTTACTGATATGACAAATTCATGTCAAACGAGA 381
   |||||
67 pLysCysAlaIleLeuLysGlnThrValArgGlnIleArgGlnIleLysG 84
   |||||
382 TAATATGGCATTTTAAAGAAACAGTAACACAGATACGTCMAATTAAG 431
   |||||
84 LulGlnGlyLysThrIleSerAsnAspAspValGlnLysAlaAspVal 100
   |||||
432 AGCAAGGAAACAACTATTCATGATGATGATGATCAAAAGCCGATGTA 481
   |||||
101 SerSerThrGlyGlnGlyValIleAspLysAspSerLeuGlyProLeu 117
   |||||
482 TCTTCTACAGGAGGAGGATTATGATTAAGACTCTTAGACCGCTTTT 531
   |||||
117 ulGlnAlaLeuAspGlyPheLeuPheValIleAsnArgGlnAlaAsnI 134
   |||||
532 ACTTCAGGCATTTGGATGTTTCCATTTTGTGTAATCGAGACGGAACA 581
   |||||
134 leValPheValSerGlnAsnValThrGlnTyrLeuGlnTyrLysGln 150
   |||||
582 TTGTATTTGATCAGAAATGTCACACAAATACCTGCAATATTAAGCAAG 631
   |||||
151 AspleuValAsnThrSerValTyrAsnIleLeuHisGlnGluAspArgly 167
   |||||
632 GACCTGGTTAACACAAAGTGTTCACAAATCTTACATGAAGAGCAAGAA 681
   |||||
167 saSPheLeuLysAsnLeuProLysSerThrValAsnGlyValSerTrp 184
   |||||
682 GGATTTTCTTAAGATTTTACCAAAATCTACAGTTAATGAGATTCCCGA 731
   |||||
184 hrAsnGluProGlnArgGlnLysSerHisThrPheAsnGlySarGlyMet 200
   |||||
732 CAATATGAGCCCAAGCAAAAGCAAAAGCATTAATTCGCCGATGTGTG 781
   |||||
201 MetLysThrProHisAspIleLeuGluAspIleAsnAlaSerProGly 217
   |||||
782 ATGAAAACCCACATGATATTCGGAAGACATTAACCCAGTCTGTAAT 831
   |||||
217 LArGlnArgTyrGlnThrMetGlnCysPheAlaLeuSerGlnProArg 234
   |||||

```

```

832 GCGCCAGATATGAAACAAATGCAAGTCTTGGCCCTGCTCAGCCAGAG 881
   |||||
234 lAeMetGlnGluGlyLysAsnLeuGlnSerCysMetIleCysValAla 250
   |||||
882 CTATGATGAGGAGGAGGAGATTTGCAATCTTGTATGATCTGTGTGCA 931
   |||||
251 ArgArgIleThrThrGlyLysArgThrPheProSerAsnProGlySerP 267
   |||||
932 CGCCGATTTACTACAGGAAAGCAATTCATCAAAACCTTGAGACCTT 981
   |||||
267 eIleThrArgHisAspLeuSerGlyValIleAsnIleAspThrAsn 284
   |||||
982 TATTACCAACATCATCTTTCAGGAAAGTTGTCATATATGATCAATAT 1031
   |||||
284 erLeuArgSerSerMetArgProGlyPheGlnAspIleIleArgArgCys 300
   |||||
1032 CACTGAGATCTCCATGAGGCTGCTTGAACATATATATCCGAAGTGT 1081
   |||||
301 lIleGlnArgPhePheSerLeuAsnAspGlyGlnSerTrpSerGlnLys 317
   |||||
1082 ATTCAGAGATTTTATGCTTAATGATGGCAGTCATGATGTCGGAAGAC 1131
   |||||
317 gHisTyrGlnGlu.....AlaTyrL 324
   |||||
1132 TCATATACAAAGAACTTACCACTGATGGGATATTTTCCCAACAGCTTAT 1181
   |||||
324 euAsnGlnHisAlaGlnThrProValTyrArgPheSerLeuAlaAspGly 340
   |||||
1182 TTAATGGCATGCGAAGAAACCCAGTATATGATCTGCTGGCGATGAGA 1231
   |||||
341 ThrIleValThrAlaGlnThrLysSerLysLeuPheArgAsnProVal 357
   |||||
1232 ACTATAGTACTGCGACAGACAAAAGCAAACTTCCGAAATCCTGTGAC 1281
   |||||
357 rAsnAspArgHisGlyPheValSerThrHisPheLeuGlnArgGlnAla 374
   |||||
1282 AAATGATCGACATGCTTTGTCTCAACCCACTTCTTCAGAGAAACAGA 1331
   |||||
374 snGlyTyrArgProAsnProAsnProValGlnGlnGlyIleArgPro 390
   |||||
1332 ATGATATATGACCAACCAACCAATCTGTTGACAAAGGATTAAGCACCT 1381
   |||||
391 MetAlaGlyCysAsnSerSerValGlyLysMetSerMetSerProAsn 407
   |||||
1382 ATGGCTGATGCAACAGTGTGTAAGCGGCAATGATGTCGCAAAACCA 1431
   |||||
407 nGlyLeuGlnMetProSerSerArgAlaTyrGlyLeuAlaAspProSer 424
   |||||
1432 AGGCTTACAGATGCGGAGACAGAGGCTATGCTTGGACAGCCCTAGCA 1481
   |||||
424 hrThrGlyGlnMetSerGlyAlaArgTyrGlyLysSerSerAsnIleAla 440
   |||||
1482 CCACAGGCGAGATGAGTGGAGCTGATGGGGGTTCCAGTAACATAGCT 1531
   |||||
441 SerLeuThrProGlyProGlyMetGlnSerProSerSerLysGlnAsn 457
   |||||
1532 TCATTGACCCCTGGGCCAGGATGCAATCAACCATCTTCAACCAAGAA 1581
   |||||
457 nasTyrGlyLeuAsnMetSerSerProProHisGlySerProGlyLeu 474
   |||||
1582 CAATATAGGCTCAACATGATGATAGCCGCCCAACAGGAGATCTGTGTT 1631
   |||||
474 lAProAsnGlnGlnAsnIleMetIleSerProArgAsnArgGlySerPro 490
   |||||
1632 CCCCAAAACCGACAGAAATATCATGATTTCTCCTCTAATCGGGAGATCA 1681
   |||||
491 LysIleLeuSerHisGlnPheSerProValAlaGlyValHisSerPhe 507
   |||||
1682 AAGTATGCTTCACTCAATGATTTTCTGTTGCAAGTGTGATCTCCAT 1731
   |||||
507 lAlaSerSerGlyAsnThrGlyAsnHisSerPheSerSerSerSerLeus 524
   |||||
1732 GGCATCTTTCGCAATACCTGGAAACCAACAGCTTTTCCAGCAGCTCTCTCA 1781
   |||||

```

524 eRLAlaLeuclnAlaIleSerGluGlyValGlyThrSerLeuLeuSerThr 540  
 1782 TGCCCTGCAAGCCATGAGTGGGGGACCTCCCTTTATCTACT 1831  
 541 LeuSerSerProGlyProLysLeuAspAsnSerProAsnMetAsnIleTh 557  
 1832 CTGTCTATCCAGAGCCCAATTTGGATTAACTCTCCCAATATGATATTAC 1881  
 557 rGlnProSerLysValSerAsnGlnAspSerLysSerProLeuGlyPheT 574  
 1882 CCAACCAAGTAAGTAAGCAATCAGATTCCAAGAGTCTCTGGGCTTTT 1931  
 574 yrcYAspGlnAsnProValGluSerSerMetCysGlnSerAsnSerArg 590  
 1932 ATTGGACACCAAAATCCAGTGGAGAGTTCAATGTCAGTCAAAATACAGA 1981  
 591 AspHisLeuSerAspLysGluSerLysGluSerSerValGluGlyLacI 607  
 1982 GATCACCTCAGTGACAAAGAAAGTAAGAGAGCAGTGTGAGGGGCGACA 2031  
 607 uAsnGlnArgGlyProLeuGluSerLysGlyHisLysLysLeuLeuGln 624  
 2032 GAATCAAAAGGGGCTTTGGAAAGCAAGGTCTATAAAATTAATCTCAGT 2081  
 624 euLeuThrCysSerSerAspAspArgGlyHisSerSerLeuThrAsnSer 640  
 2082 TACTTACCTGTTCTTCTGATGACCGGGGTCATTCCTCCCTTGACCACTCC 2131  
 641 ProLeuAspSerSerCysLysGluSerSerValSerValThrSerProse 657  
 2132 CCCCTGATTTCAAGTTGTAAGAAATCTTCTGTAGTGTACCAAGCCCTTC 2181  
 657 rGlyValSerSerSerThrSerGlyValSerSerThrSerAsnMetH 674  
 2182 TGGAGTCTCTCTCTACATCTGAGAGAGTATCCCTCATCATCAATATATGC 2231  
 674 IsGlySerLeuLeuGlnGlyHisArgIleLeuHisLysLeuLeuGln 690  
 2232 ATGGGCTACTGTTACAAAGAAAGCACCGGATTTTGACAAAGTTGCGCAG 2281  
 691 AsnGlyAsnSerProAlaGluValAlaLysIleThrAlaGlnAlaThrG 707  
 2282 AATGGGAATTCACACACTGAGCTAGCCAAAGATTACTGCAGAAAGCCACTGG 2331  
 707 yLysAspThrSerSerIleThrSerCysGlyAspGlyAsnValIleLysG 724  
 2332 GAAAGACACACAGATATAACTTCTTGCGGGACGGAATGTTCACAAC 2381  
 724 InGluGlnLeuSerProLysLysGlyAsnAsnAlaLeuLeuArgIle 740  
 2382 AGGAGCAGCTAAGTCTTAAGAAAGAGAAATATGCACTTCTTAATATAC 2431  
 741 LeuLeuAspArgAspAspProSerAspAlaLeuSerLysGluLeuGlnP 757  
 2432 CTGCTGGACAGGAGATCTTACTGATGCACCTCTTAAGAAACTACAGCC 2481  
 757 oGlnValGluGlyValAspAsnLysMetSerGlnCysThrSerSerThr 774  
 2482 CCAAGTGGAAAGAGTGAATTAATTAATGATCAGTCCACAGCTCCACACA 2531  
 774 LeProSerSerSerGlnGlyLysAspProLysIleLysThrGluThrSer 790  
 2532 TTCTTAAGCTCAAGTCAAGAAAGACCCCTTAATTAAGACAGAGCAACT 2581  
 791 GluGluGlySerGlyAspLeuAspAsnLeuAspAlaIleLeuGlyAspLe 807  
 807 uThrSerSerAspPheTyAsnAsnSerIleSerSerAsnGlySerHisL 824  
 2632 GACTAGTCTGACTTTTACAAATATTCATTCCTCAATGCTAGTCATC 2681

824 euGlyThrLysGlnGlnValPheGlnGlyThrAsnSerLeuGlyLeuLys 840  
 2682 TGGGACTAAGCAACAGGTGTTCAAGGAACTAATCTCTGGGTTTGAAA 2731  
 841 SerSerGlnSerValGlnSerIleArgProProTyAsnArgAlaValaSe 857  
 2732 AGTTTCACATCTGTGACATCTATTCGTCTCCATATATACCGAGCACTGTC 2781  
 857 rLeuAspSerProValSerValGlySerSerProProValLysAsnIleS 874  
 2782 TCTGATACCCCTGTTTCTGTGGCTCAAGTCTCCAGTAATAAAATATACA 2831  
 874 eRLAAlaPheProMetLeuProLysGlnProMetLeuGlyLysAsnProArg 890  
 2832 GTGCTTTCCCATGTTTACCAAGCAACCCATGTGTGGTGGAAATCCACA 2881  
 891 MetMetAspSerGlnGluAsnTyrgLysSerSerMetGlyLysProAsnR 907  
 2882 ATGATGATAGTTCAGGAAATATATGCTCAAGTATG. .... 2917  
 907 gAsnValThrValThrGlnThrProSerSerGlyAspTrpGlyLeuPro 924  
 2918 .....GAGACTGGGCTTACCA 2936  
 924 snSerLysAlaGlyArgMetGluProMetAsnSerAsnSerMetGlyArg 940  
 2937 ACTCAAAAGCCGCGACAAATGAACTATGAAATTCAAACTCCATGGGACAA 2986  
 941 ProGlyGlyAspTyraAsnThrSerLeuProArgProAlaLeuGlyLysE 957  
 2987 CCAAGAGGAGATTATATATCTTATACCAAGACCTGCACTGGTGCTC 3036  
 957 rIleProThrLeuProLeuArgSerAsnSerIleProGlyAlaArgProV 974  
 3037 TATTCACACATTCCTCTGCTGCTAATAGCATACAGGTGGAGACAGCAG 3086  
 974 AlLeuGlnGlnGlnGlnMetLeuGlnMetArgProGlyLulIlePro 990  
 3087 TATTCACACAGCAGCAGCATGCTTCAATAGAGGCTGTGTAATATCCCC 3136  
 991 MetGlyMetGlyAlaAsnProTyrgLysGlnAlaAlaSerAsnGlnLe 1007  
 3137 ATGGGAATGGGGCTATATCCCTATGGCCAAAGCAAGCACTTACCAACT 3186  
 1007 uGlySerTrpProAspGlyMetLeuSerMetGluGlnValSerHisGlyT 1024  
 3187 GGGTTCCTGGCCCGAATGGCATGTGTCCATGGAACAAGTTCTCATGGCA 3236  
 1024 hrcGlnAsnArgProLeuLeuArgAsnSerLeuAspLeuValGlyPro 1040  
 3237 CTCAAATATGGCCCTCTTCTTAAGAAATTCCTGATGATCTTGTGGCCA 3286  
 1041 ProSerAsnLeuGluGlyGlnSerAspGluArgAlaLeuLeuAspGlnLe 1057  
 3287 CCTTCCAACTGGAAGGCCAGAGTCAAGAAAGCAATTAATTTGACACAGCT 3336  
 1057 uHisThrLeuLeuSerAsnThrAspAlaThrGlyLeuGluGluIleAspA 1074  
 3337 GCACACTTCTTCAGAACACACAGATGCCACAGCGCTGGAAAGAAATTGACA 3386  
 1074 rGAlaLeuGlyIleProGluLeuValAsnGlnGlyGlnAlaLeuLupro 1090  
 3387 GAGCTTTGGGCATTCCTGAACCTTGCAACACAGGAGCAGCATTAAGGCC 3436  
 1091 LysGlnAspAlaPheGlnGlyGlnGluAlaAlaValMetMetAspGlyL 1107  
 3437 AAACAGGATCTTTCCAAAGGCCAAAGAGCAGCAGTAATGATGATCAGAGA 3486  
 1107 sAlaGlyLeuTyrgLysGlnThrTyrgProAlaGlnGlyProprometGln 1124  
 3487 GGCAGGATATATGACAGACATACCAAGCAGCAGGAGGCTCCCAATGCAAG 3536  
 1124 LysGlyPheHisLeuGlnGlyGlnSerProSerPheAsnSerMetMetAsn 1140

```

3537 GAGCTTTCATCTTCAGGACATCACCATCTTTAACTCATATGAT 3586
1141 GlnMetAsnGlnGlnGlnAsnProLeuGlnGlnMetHisProArgAl 1157
3587 CAGATGAACGACGAAAGCAATTTTCTCTCCAGAAATGACCCAGCAGC 3636
1157 aAsnIleMetArgProArgThrAsnThrProLysGlnLeuArgMetGln 1174
3637 CAACATCATGAGACCCCGGACAAACACCCCAACCACTTAAGAAATGACG 3686
1174 euGlnIleArgLeuGlnGlnGlnIlePheLeuAsnGlnSerArgGlnAla 1190
3687 TTCAGCAGAGGCTCGAGGCGCAGAGTTTGGAAATCAGAGCGCAGAGCA 3736
1191 LeuGlnLeuLysMetGluAsnProThrAlaGlyAlaAlaValMetAr 1207
3737 CTTGAATTGAAATGAAACCCCTACTGCTGTGCTGCTGCGTGAATGAG 3786
1207 gProMetMetGlnProGlnGlnGlnIlePheLeuAsnAlaGlnMetValAla 1224
3787 GCGTATGATGAGCCCGACAGGGGTTTCTTAATGCTCAATVGGTCGCC 3836
1224 IlnArgSerArgGlnLeuLeuSerHisHisPheArgGlnIleArgValAla 1240
3837 AACGCACAGAGAGCTGCTAAGTCATCATCTCCACAAACAGAGGGTGCT 3886
1241 MetMetMet.....GlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1253
3887 ATGATGATGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 3936
1253 nGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1270
3937 GCAGCAACAGCAACAGCAACAGCAACAGCAACAGCAACAGCAACAGC 3986
1270 IlnPheSerProProAsnValThrAlaSerProSerMetAspGlyLeu 1286
3987 CCTTCAGCCGACCTCTAATGTGACTGTTCCCGCAGCATGATGGGCTT 4036
1287 LeuAlaGlyProThrMetProGlnAlaProProGlnIlePheProtyrG 1303
4037 TTGGCAGAGCCCAATGCCCAAGCTCTCCGCAACAGTTTCCATATCA 4086
1303 nProAsnTyrgLysMetGlyGlnIleProAspProAlaPheGlyArgVal 1320
4087 ACCAAATATGAAATGGACAAACAGATCCAGCTTGTGTGAGTGT 4136
1320 eRserProProAsnAlaMetMetSerSerArgMetGlyProSerGlnAsn 1336
4137 CTAGTCCGCCCAATGCAATGATGTGCTCAAGATGGGTCCCTCCAGAT 4186
1337 ProMetMetGlnHisProGlnAlaAlaSerIleTyrgLysSerSerGlu 1353
4187 CCCATGATGCAACACCCGACAGCTGCATCCATCATCTAGTCTCAGAA 4236
1353 tLysGlyTyrProSerGlyAsnLeuAlaArgAsnSerSerPheSerGln 1370
4237 GAAGGCTGGCCATCAGAAATTTGGCCAGGACAGCTCTTTTCCAGC 4286
1370 IlnGlnPheAlaHisGlnIleAsnProAlaValTyrSerMetValHisMet 1386
4287 AGCAGTTTGCCACACAGGGAATCCTGCAGTATAGTATGTGCACATG 4336
1387 AsnGlySerSerGlyHisMetGlyIleMetAsnMetAsnProMetProme 1403
4337 AATGGCAGCACTGTGCATGTGAGAGATGAACATGAACCCCATGCCAT 4386
1403 tSerGlyMetProMetGlyProAspGlnLysTyrcys 1415
4387 GTCTGGCATGCTATGGGTCTGTGATCAGAAATACTGC 4423
seq_name: /cgn2_6/prodata/1/pna/US0966.COMB.seq:US-09-698-014-5135

```

```

seq_documentation_block:
; Sequence 5135, Application US/09698014
; GENERAL INFORMATION:
; APPLICANT: Geating, David P.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Kingsbury, Gillian A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.2014-001
; CURRENT APPLICATION NUMBER: US/09/698, 014
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: 60/162,363
; NUMBER OF SEQ ID NOS: 6098
; SOFTWARE: FASTSeq for Windows Version 4.0
; SEQ ID NO 5135
; LENGTH: 7888
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-698-014-5135

alignment_scores:
Quality: 7233.50 Length: 1429
Ratio: 5.174 Gaps: 3
Percent Similarity: 97.831 Percent Identity: 97.621

alignment_block:
US-09-041-994-2 x US-09-698-014-5135 ..
Align seg 1/1 to: US-09-698-014-5135 from: 1 to: 7888

1 MetSerGlyLeuGlyGluAsnLeuAspProLeuAlaSerAspSerArgly 17
182 ATGATGATGATGAGAAAGAACTTGATCCAGTGGCCAGTATTCAGCAAA 231
17 sArgLysLeuProCysAspThrProGlyGlnGlyLeuThrCysSerGly 34
232 AGCAAAATGCCATGTGATCTCCAGCAAGGCTTACCTGACGTGAGTG 281
34 IulysArgArgArgGluGlnGlnLeuSerLysTyrIleGluGluLeuAlaGlu 50
282 AAAAAGCAGAGCGGAGCAGCAAGTAATATATGAAGATGCGTGCAG 331
51 LeuIleSerAlaAsnLeuSerAspIleAspAsnPheAsnValLysProAs 67
332 CTGATATCTGCCAATCTTAGATATGACAAATTCATATGCAATGCAACCA 381
67 PLYSCysAlaIleLeuLysGluThrValArgGlnIleArgGlnIleLysG 84
382 TAAATGTGCGATTTTAAAGCAACGTAAGACAGATGATGTTCAAAAAGCCGATG 431
84 IuGlnGlyLysThrIleSerAsnAspAspValGlnLysAlaAspVal 100
432 AGCAAGGAAAAAATATTCATGATGATGATGATGTTCAAAAAGCCGATG 481
101 SerSerThrGlyGlnGlyValIleAspLysAspSerLeuGlyProLeu 117
482 TCTTCTACAGGCGAGGAGTATGATTAAGACATCTTAGACCGCTTTT 531
117 uLeuGlnAlaLeuAspGlyPheLeuPheValAlaAsnArgGlnAlaAsn 134
532 ACTTCAGCATTTGGATGTTTCTATTTTGTGTGATGATGATGATGATGATG 581
134 IeValPheValSerGluAsnValThrGlnTyrLeuGlnTyrLysGlnGlu 150
582 TTGATATTGTATCAGAAATGTCACACATACCTGCAATATATTAACAGAG 631
151 AspLeuValAsnThrSerValTyrAsnIleLeuHisGluGluAspArgly 167
632 GACCTGTGTTAACACAAAGTGTATACAAATATCTTACATGTAAGAACAGAA 681
167 sAspPheLeuLysAsnLeuProLysSerThrValAsnGlyValSerTrp 184

```

```

|||||
682 GGATTTTCTTAAGAAATTTCACAAAATCTACAGTTAATGAGATTTCCTCGGA 731
|||||
184 hAsnGluProGlnArgGlnLysSerHisThrPheAsnCysArgMetLeu 200
|||||
732 CAAATGAGACCCCAAGACAAAAGCCATATCTTAAATGCGGTATGTTG 781
|||||
201 MetLysThrProHisAspIleLeuGlnAspIleAsnAlaSerProGluMet 217
|||||
782 ATGAAAAACCACTGATGATATCTGGAAAGCATTAACGCCAGTCTCTGAAT 831
|||||
217 tArgGlnArgTyrGluThrMetGlnCysPheAlaLeuSerGlnProArg 234
|||||
832 GCGCCAGAGATATGAACAATGACAGTCTTTCCTGCTCAGCCACGAG 881
|||||
234 lMetMetGlnGlnGlyLysAspLeuGlnSerCysMetIleCysValAla 250
|||||
882 CTATGATGAGAGAGGGAAGATTTCATCTTGATGATCTGTGTGGCA 931
|||||
251 ArgArgIleThrThrGlyGluArgThrPheProSerAsnProGluSerPh 267
|||||
932 CCGCCGATTACTACAGAGAAAAGACATTTCCATTCAAACCTGAGAGCTT 981
|||||
267 eIleThrArgHisAspLeuSerGlyLysValAlaAsnIleAspThrAsn 284
|||||
982 TATTACCAACATGATCTTTCAGAGAAAGTTGTCAATATGATACAAAT 1031
|||||
284 eLeuArgSerSerMetArgProGlyPheGluAspIleIleArgArgCys 300
|||||
1032 CACTGAGATCTTCATGAGGCGCTGGCTTGAAGATTAATCCGAAGCTGT 1081
|||||
301 lIleGlnArgPhePheSerLeuAsnAspGlyGlnSerTrpSerGlnLys 317
|||||
1082 ATTCAAGAGATTTTAAATGATGAGGTCATGGTCCCAAAAG 1131
|||||
317 gHisTyrGlnGlu.....AlaTyrL 324
|||||
1132 TCACATACAAAGAAATTACCACTGATGGGATATTTCGCCAACAGCTTATC 1181
|||||
324 euAsnGlyHisAlaGlnThrProValTyrArgPheSerLeuAlaAspGly 340
|||||
1182 TTAATGGCCATGCAGAAACCCAGTATTCGATTCGTTGGCTGATGGA 1231
|||||
341 ThrIleValThrAlaGlnThrLysSerLysLeuPheArgAsnProValh 357
|||||
1232 ACTATAGTACTGCACAGACAAAAGCAACCTTCGAAATCCGTAAAC 1281
|||||
357 rAsnAspArgHisGlyPheValSerThrHisPheLeuGlnArgGluGln 374
|||||
1282 AATATGATGCATGGCTTGTCTCAACCCACTTCCTTCAGAGAGAAACGA 1331
|||||
374 snGlyTyrArgProAsnProAsnProValGlyGlnGlyIleArgProPro 390
|||||
1332 ATGGATATAGACCAAAACCAATTCCTGTGGACAAGGATTAGACACCT 1381
|||||
391 MetAlaGlyCysAsnSerSerValGlyGlyMetSerMetSerProAsnAl 407
|||||
1382 ATGGCTGGATGCAACAGTTCGGTAGGCGCATGAGATATCGCCAAACCA 1431
|||||
407 nGlyLeuGlnMetProSerSerArgAlaTyrGlyLeuAlaAspProSer 424
|||||
1432 AGGCTTACAGATGCCGAGACGAGGCTTATGGCTTGGCAGACCCCTAGA 1481
|||||
424 hThrGlyGlnMetSerGlyAlaArgTyrGlyLysSerSerAsnIleAla 440
|||||
1482 CCACAGGGCAGATGAGTGGAGCTAGGTATGGGGGTCCAGTAACATAGCT 1531
|||||
441 SerLeuThrProGlyProGlyMetGlnSerProSerSerTyrGlnAsnAs 457
|||||
1532 TCATTGACCCCTGGGCGCAGCATGCAATCACATCTTCTTACCAGAAACA 1581
|||||
457 nAsnTyrGlyLeuAsnMetSerSerProHisGlySerProGlyLeuAla 474
|||||

```

```

1582 CAACCTATAGCTCAACATGAGTAGCCCCCACATGGGAGATCCTCGTCTTC 1631
|||||
474 lAsProAsnGlnGlnAsnIleMetIleSerProArgAsnArgLysPro 490
|||||
1632 CCCCAAAACGACGAAATATCATGATTTCTCTCGTAACTGTGGAGTCCA 1681
|||||
491 LysIleAlaSerHisGlnPheSerProValAlaGlyValHisSerProMet 507
|||||
1682 AAGATAGCCTTCACATCAGTTTCTCTGTGGACGGTGTCTCCAT 1731
|||||
507 tAlaSerSerGlyAsnThrGlyAsnHisSerPheSerSerSerSerLeu 524
|||||
1732 GGCATTTCTGGCAATACCTGGGACACAGACCTTTTCACAGACTCTCTCA 1781
|||||
524 eAlaLeuGlnAlaIleSerGlnGlyValGlyThrSerLeuSerThr 540
|||||
1782 GTGCCCTGCAGCCATCAGTGAAGGTGTGGGGACTTCCTTTATCTACT 1831
|||||
541 LeuSerSerProGlyProLysLeuAspAsnSerProAsnMetAsnIleThr 557
|||||
1832 CTGTCAATCACAGGCCCAAAATGATTACTCTCCCAATATGAAATATAC 1881
|||||
557 rGlnProSerLysValSerAsnGlnAspSerLysSerProLeuGlyPhe 574
|||||
1882 CCAACCAATTAAGTAAGCAATCAGGATTCCAAGACTCCTTGCGCTTTT 1931
|||||
574 yrCysAspGlnAsnProValGluSerSerMetCysGlnSerAsnArg 590
|||||
1932 ATTGCAGCCAAATACAGTGGAGAGTTCAATGTCAGTCAATATGACAGA 1981
|||||
591 AspHisLeuSerAspLysGluSerLysGluSerSerValGlnGlyAlaGln 607
|||||
1982 GATCACCTCAGTGACAAAGAAAGTAAGGAGACAGAGTTGAGGGGACAA 2031
|||||
607 uAsnGlnArgGlyProLeuGluSerLysGlyHisLysLysLeuGlnL 624
|||||
2032 GAATCAAAAGGCTCTTTGGAAAGCAAAAGCTATATAAAATATACGCACT 2081
|||||
2082 TACTTACCTGTTCTTCTGATGACCGGGGTCAATTCCTCTGACCAACTCC 2131
|||||
641 ProLeuAspSerSerCysLysGluSerSerValSerValThrSerPro 657
|||||
2132 CCCCTAGATTCAAGTGTAAAGAAATCTTGTAGTGTACACAGCCCCCTC 2181
|||||
657 rGlyValSerSerSerThrSerGlyValSerSerThrSerAsnMet 674
|||||
2182 TGGAGTCTCTCCTCTACATCTGGAGAGTATCTTACATCCAAATATGC 2231
|||||
674 lAsGlySerLeuLeuGlnGlnLysHisArgIleLeuHisLysLeuGln 690
|||||
2232 ATGGGTCACCTGTTACAAAGAAAGCACCGGATTTTGCACAAATGGTGCAG 2281
|||||
691 AsnGlyAsnSerProAlaGluValAlaLysIleThrAlaGlnAlaThrGln 707
|||||
2282 AATGGAAATTCACCAAGCTGAGTACGCCAAGATTACTGCGAAAGCCACTGG 2331
|||||
707 yLysAspThrSerSerIleThrSerCysGlyAspGlyAsnValAlaLysG 724
|||||
2332 GAAAGACACACAGAGATATTAACCTTTGTGGGACGGAATGTGTTCACAC 2381
|||||
724 lGlnGlnLeuSerProLysLysGlyAsnAsnAlaLeuLeuArgTyr 740
|||||
2382 AGGAGCAGCTTAAGTCTTAAGAAAGAAAGAAATTAATGCACTTCTTAGATAC 2431
|||||
741 LeuLeuAspArgAspAspProSerAspAlaLeuSerLysGluLeuGlnPro 757
|||||
2432 CTGCTGACAGGAGATGATCTAGTGTGACACTCTTAAGAAACTCAGAGCC 2481
|||||
757 oGlnValGlnGlyValAspAsnLysMetSerGlnCysThrSerSerThrI 774
|||||
2482 CCAAGTGAAGAGAGTGCAATTAATAATGATGATGTCAGACACACTCCACCA 2531
|||||

```

774 LeprosinSerSerGlnGlnLysAspProLysIleLysThrGluThrSer 790  
|||||  
2532 TTCTAGCTCAAGTCAAGAGAAAGCCCTAAATAATTAAGACAGACCAAGT 2581  
791 GluGlnGlySerGlyAspLeuAspAsnLeuAspAlaIleLeuGlyAspLe 807  
|||||  
2582 GAAGAGGATGTGAGACTTGATGATATCTATGATCTTCTGGTGATCT 2631  
807 uThrSerSerAspPheTyrAsnAsnSerIleSerSerAsnGlySerHisL 824  
|||||  
2632 GACTAGTCTCACTTTACAAATTAATCCATATCCCTCAAAATGATGATC 2681  
824 euGlyThrLysGlnGlnValPheGlnGlyThrAsnSerLeuGlyLeuAs 840  
|||||  
2682 TGGGAGCAACCAAGAGTGTTCAGAGAACTAATCTCTGGGTTTGAA 2731  
841 SerSerGlnSerValGlnSerIleArgProProTyrAsnArgAlaValSe 857  
|||||  
2732 AGTTCACAGCTGTGCAAGTATTCGTCTCCATATAACGACGACGTCTC 2781  
857 rLeuAspSerProValSerValGlySerSerProProValLysAsnIleS 874  
|||||  
2782 TCTGGAAGCCCTGTTCTGTCGCTCAAGTCAAGTCAAAATAATATCA 2831  
874 eAlaPheProMetLeuProLysGlnProMetLeuGlyLysAsnProArg 890  
|||||  
2832 GTGCTTTCCTCCATGTTCACCAAGCAACCCATGTGGGGAATTCAGAA 2881  
891 MetMetAspSerGlnGlnAsnTyrGlySerSerMetGlyGlyProAsnAr 907  
|||||  
2882 ATGATGATGATGCTCAGAAATTAATGCGTCAAGTATG..... 2917  
907 gAsnValThrValThrGlnThrProSerSerGlyAspTyrPglLeuProA 924  
|||||  
2918 .....GGAGACTGGGGCTTACCAA 2936  
924 snSerLysAlaGlyArgMetGluProMetAsnSerAsnSerMetGlyArg 940  
|||||  
2937 ACTCAAAAGCGCGCAAGATGAGACCTTAAGAAATTAACCTCATGGGAAGA 2986  
941 ProGlyGlyAspTyrAsnThrSerLeuProArgProAlaLeuGlyLysE 957  
|||||  
2987 CCAGAGAGAGATTAATTAATCTTCTTACCCAGACTGCTGGGTGGCTC 3036  
957 rIleProThrLeuProLeuArgSerAsnSerIleProGlyAlaArgProV 974  
|||||  
3037 TATTCACCAATTCCTCTCGGTCTAATAGCATACAGAGTCCGAGACCAG 3086  
974 alleuGlnGlnGlnGlnMetLeuGlnMetLeuArgProGlyGluIlePro 990  
|||||  
3087 TATTCACCAACAGCAGCAGATGCTTCAAAATGAGGCGGTGAATATCCCT 3136  
991 MetGlyMetGlyAlaAsnProTyrGlyAlaIleAlaIleSerAsnGlnLe 1007  
|||||  
3137 ATGGGAATGGGGGCTAATCCCTATGCGCAAGCAGACAGCATTAACCAACT 3186  
1007 uGlySerTyrProAspGlyMetLeuSerMetGlyGlnValSerHisGlyT 1024  
|||||  
3187 GGGTTCCTGGCCCGATGATGTTGCCATGGAACAAGTTCTCATGGCA 3236  
1024 hrGlnAsnArgProLeuLeuArgAsnSerLeuAspAspLeuValGlyPro 1040  
|||||  
3237 CTCAAAATAGCCCTCTCTAGGAATTCCTCGGATGATCTTGTGGGCGCA 3286  
1041 ProSerAsnLeuGlnGlnSerAspGlnArgAlaLeuLeuAspGlnLe 1057  
|||||  
3287 CTTTCCAACTCGGAAGCGCAGAGTGAAGAAAGACATTATTTGACCACT 3336  
1057 uHisThrLeuLeuSerAsnThrAspAlaThrGlyLeuGlnGluIleAspA 1074  
|||||  
3337 GCACAGCTCTTCTCAGCAACACAGATGCCACAGGCGCTGGAAGAAATTTGACA 3386

1074 rGAlaLeuGlyIleProGluLeuValAsnGlnGlyGlnAlaLeuGluPro 1090  
|||||  
3387 GAGCTTTGGGCAATCTGAATCTGCAATCAAGGACAGCATTAAGAGCC 3436  
1091 LysGlnAspAlaPheGlnGlnGlnGlnAlaIleAlaValMetLeuAspGlnLy 1107  
|||||  
3437 AAACAGATGCTTTCACAAAGCCAGAAAGACAGCAGTAATGATGATCAGAA 3486  
1107 sAlaGlyLeuTyrGlyGlnThrTyrProAlaGlnGlyProProMetGlnG 1124  
|||||  
3487 GGCAGGATTAATATGACAGACATACCCAGCAGAGGGGCTCAATGCAAG 3536  
1124 LyGlyPheHisLeuGlnGlnGlnSerProSerPheAsnSerMetLeuAsn 1140  
|||||  
3537 GAGCTTTCACTTCAGAGCAATCACCATCTTTTAATCTATGATGAT 3586  
1141 GlnMetAsnGlnGlnGlnLysAsnPheProLeuGlnGlnMetHisProArgAl 1157  
|||||  
3587 CAGATGAACCAAGCAAGCAATTTCTCTCCAGAGAAATGACCAAGCAGCAG 3636  
1157 aAsnIleMetArgProArgThrAsnThrProLysGlnLeuArgMetGlnL 1174  
|||||  
3637 CAACATCATGAGACCCCGCAAAACCCCAACCACTTGAATGACAGC 3686  
1174 euGlnGlnArgLeuGlnGlnGlnPheLeuAsnGlnSerArgGlnAla 1190  
|||||  
3687 TTCAGCAGAGGCTGAGGGCCAGCAGTTTGAATCAAGCCGACAGCA 3736  
1191 LeuGlyLeuLysMetGlyAsnProThrAlaGlyGlyAlaIleAlaValMetAr 1207  
|||||  
3737 CTTGATTTGAATGGAATGGAACCCCTACTCTGCTGCTGGGCGATGAG 3786  
1207 gProMetMetGlnProGlnGlnGlyPheLeuAsnAlaGlnMetValAlaG 1224  
|||||  
3787 GCTATGATGACAGCCCGCAGAGGTTTCTTAATGCTCAATATGTCGGCC 3836  
1224 lnaArgSerArgGluLeuLeuSerHisHisPheArgGlnGlnArgValAla 1240  
|||||  
3837 AACGCAGCAGAGCTGCTAAGTCATCACTTCCAGCAACAGAGGTGCT 3886  
1241 MetMetMet.....GlnGlnGlnGlnGlnGlnGlnGlnGln 1253  
|||||  
3887 ATGATGATGACGACGACGACGACGACGACGACGACGACGACGACGACGA 3936  
1253 nGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1270  
|||||  
3937 GCAGCAACAGCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 3986  
1270 lAphSerProProProAsnValThrAlaSerProSerMetAspGlyLeu 1286  
|||||  
3987 CTTTACGCCACCTCTAATGTGACTGCTTCCCCAGCATGATGGGCTT 4036  
1287 LeuAlaGlyProThrMetProGlnAlaProProGlnGlnPheProTyrGly 1303  
|||||  
4037 TTGGCAGAGACCCACATGCCAAGCTCTCTCCGAAACAGTTTCCATATCA 4086  
1303 nProAsnTyrGlyMetGlyGlnGlnProAspProAlaPheGlyArgValS 1320  
|||||  
4087 ACCAAATATGATGAGTGGACAACAACAGATCCAGCTTGTGCGAGTGT 4136  
1320 eRserProProAsnAlaMetMetSerSerArgMetGlyProSerGlnAsn 1336  
|||||  
4137 CTAGTCCCTCCCAATGCAATGATGTCGTCMAAAGTGGGCTCTCCAGAT 4186  
1337 ProMetMetGlnHisProGlnAlaIleSerIleTyrGlnSerSerGlyMe 1353  
|||||  
4187 CCCATGATGACACACCCGACAGCTGCAATCATCATAGTCTTAGAAT 4236  
1353 tLysGlyTyrProSerGlyAsnLeuAlaArgAsnSerSerPheSerGlnG 1370  
|||||  
4237 GAAGGGCGGCGCATCAGGAATTTGGCAGAGAACAGCTTTTCCAGCAG 4286  
1370 lGlnPheAlaHisGlnGlnLysAsnProAlaValTyrSerMetValHisMet 1386

|||||  
4287 ACAGATTGCCACCGAGGGAATCTGCAGTGTATGATGGCGACAG 4336  
1387 AaNgLySeSerGlyHisMetGlyGlnMetAsnMetAsnProMetPro 1403  
4337 AATGCAGCAGAGTGTGCATGCGACAGATGACATGACCCCATGCCCAT 4386  
1403 tSerGlyMetProMetGlyProAspGlnLysTyrCys 1415  
4387 GTCGTGCATGCCATGTGCTCGATGACGAATATCTGC 4423

seq\_name: /cgn2\_6/ptodata/1/pna/US097A\_COMB.seq:US-09-716-475-7321

seq\_documentation\_block:  
: Sequence 7321, Application US/09716475  
: GENERAL INFORMATION:  
: APPLICANT: Shyjan, Andrew W.  
: APPLICANT: Macbeth, Kyle J.  
: APPLICANT: Vasicek, Thomas  
: TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES  
: FILE REFERENCE: 1600.2037-001  
: CURRENT APPLICATION NUMBER: US/09/716.475  
: PRIOR FILING DATE: 2000-11-20  
: PRIOR APPLICATION NUMBER: 60/166,506  
: NUMBER OF SEQ ID NOS: 8194  
: SOFTWARE: FastSeq for Windows Version 4.0  
: SEQ ID NO 7321  
: LENGTH: 7888  
: TYPE: DNA  
: ORGANISM: Homo sapiens  
US-09-716-475-7321

alignment\_scores:  
Quality: 7233.50 Length: 1429  
Ratio: 5.174 Gaps: 3  
Percent Similarity: 97.831 Percent Identity: 97.621

alignment\_block:  
US-09-041-994-2 x US-09-716-475-7321 ..

Align seg 1/1 to: US-09-716-475-7321 from: 1 to: 7888

1 MetSerGlyLeuGlyLysAsnLeuAspProLeuAlaSerAspSerArgly 17  
182 ATGAGTGGATGTAGAGAAAACCTTGATCCACGTGCCAGTGATTCACGAGA 231  
17 sArgLySeuProCysAspThrProGlyGlnGlyLeuThrCysSerGly 34  
232 ACGCAAAATGGCCATGTGATCTCCAGACAGGCTTACCTGCAGGTGG 281  
34 LuLySArGrArGrGlnGlnLysSerLysTyrLLeGlnGlnLeuAlaGlu 50  
282 AAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 331  
51 LeuLLeSerAlaAsnLeuSerAspIleAspAsnPheAsnValLysProAs 67  
332 CTGATATCTGCCAATCTTAGTGATATTGACAAATTCATGTCAACACAGA 381  
67 pLyScySAIaIleLeuLySGluThrValArgGlnIleArgGlnIleLyG 84  
382 TAAATGTGGATTTTAAAGAAACGTAAGACAGATACGTCAAAATAAAG 431  
84 LuGlnGlyLysThrLLeSerAsnAspAspValGlnLysAlaAspVal 100  
432 ACCAAGAAACAACTATTTCATGATGATGATGATGATGATGATGATGAT 481  
101 SerSerThrGlyGlnGlyValIleAspLysAspSerLeuGlyProLeu 117  
482 TCTTCTACAGGCGAGGAGGTATTGATTAAGACTCTTAGACCCGCTTTT 531

117 uLeuGlnAlaLeuAspGlyPheLeuPheValValAsnArgGluAlaAsn 134  
532 ACTTCAGGCAATGTGATGTTCTCTATTGTGGTGAATCAGACGAAACA 581  
134 LeValPheValSerGlyAsnValThrGlnTyrLeuGlnTyrLysGln 150  
582 TTGTATTGTATCAGAAATGTCCACAAATACCTGCATATATATAGCAAG 631  
151 AspLeuValAsnThrSerValTyrAsnIleLeuHisGlyLysAspArg 167  
632 GACCTGGTTAACCAACAGTGTTCATATCTTCATGAAAGAACGAGAA 681  
167 sAspPheLeuLysAsnLeuProLysSerThrValAsnGlyValSerTyr 184  
682 GCAATTTCTTAAGAAATTTACAAAATCTACAGTTAATGAGGTTCTCTGA 731  
184 hPAsnGluProGlnArgGlnLysSerHisThrPheAsnGlyMetLeu 200  
732 CAATGAGACCCAAAGCAAAAAGCCATACATTTAATGCGCTATGTG 781  
201 MetLysThrProHisAspIleLeuGlnAspIleAsnAlaSerProGlu 217  
782 ATGAAAACACACATGATATTCTGGAAGACATTAAGCCAGTCTGAAAT 831  
217 tArgGlnArgTyrGlyThrMetGlnCysPheAlaLeuSerGlnProArg 234  
832 GCCCAGAGATATGAAACATGACAGTGTGGCCCTGCTCCAGCCAGAG 881  
234 lAeMetGluGlnGlyLysAspLeuGlnSerCysMetLLeCysValAla 250  
882 CTATGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 931  
251 ArgArgIleThrThrGlyLysValThrPheProSerAsnProGlySer 267  
932 CGCCGCAATTAATACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 981  
267 eIleThrArgHisAspLeuSerGlyLysValValAsnIleAspThrAsn 284  
982 TATTACCAACATGATCTTTCAGGAAAGGTTGCCAATATAGATACAAAT 1031  
284 eLeuArgSerSerMetArgProGlyPheGluAspIleIleArgArgCys 300  
1032 CACTGAGATCCTTCATGAGGCGCTTGGAAGATTTAATCCGAAAGGT 1081  
301 lLeGlnArgPhePheSerLeuAsnAspGlyGlnSerTyrSerGlnLys 317  
1082 ATTCAAGATTTTATGCTTAATGATGGCAGTATGATCCCAAGAAAG 1131  
317 gHisTyrGlnGlu.....AlaTyrL 324  
1132 TCACATATCAAGAAAGTTACCAAGTATGATTTTCCCAACAGCTTATC 1181  
324 euAsnGlyHisAlaGluThrProValTyrArgPheSerLeuAlaAspGly 340  
1182 TTATATGCCATGACAGAAACCCAGTATATCGATTCCTGTGGCTGATGA 1231  
341 ThrIleValThrAlaGlnThrLysSerLysLeuPheArgAsnProVal 357  
1232 ACTATAGTACAGTGCACAGCAAAAAGCAAACTTCGCCAAATCTGTAC 1281  
357 rAsnAspArgHisGlyPheValSerThrHisPheLeuGlnArgGluGln 374  
1282 AATGATGCAACATGGCTTGTCTCAACCCACCTTCAGAGAGAGACAGA 1331  
374 sGlnTyrArgProAsnProAsnProValGlnGlyLysIleArgPropo 390  
1332 ATGAGATATGACCAACCAATCTCTGTGGACAAAGGATTTAGACACCT 1381  
391 MetAlaGlyCysAsnSerSerValGlyGlyMetSerMetSerProAsn 407  
1382 ATGGCTGATGCAACAGTTCGGTAGGCGGCAATGATATGCCCAAAACA 1431  
407 nGlyLeuGlnMetProSerSerArgAlaTyrGlyLeuAlaAspProSer 424



```

|||||
1432 AGGCTTACAGATCCGACGACGAGGCGCTATGCGCTTGCGACAGCCCTAGCA 1481
424 hTThGlyLInMetSerGlyAlaArgTyrGlyLysSerSerInIleAla 440
|||||
1482 CCACAGGCGAGATGAGTGAAGTATGAGTGGGGCTCCAGTAACATAGCT 1531
441 SerLeuThProGlyProGlyMetGInSerProSerSerTyrGInAsnAs 457
|||||
1532 TCATTGACCCCTGGGCGGACGATGCATACCATCTTCTTACCAGAACAA 1581
457 nAsnTyrGlyLeuAsnMetSerSerProProhIsgLysSerProGlyLeuA 474
|||||
1582 CAACGTATAGCTCAACATGATGATAGCCCCCACATGGAGAGTCCGTGCTTG 1631
474 lAProAsnGInGInAsnIleMetIleSerProArgAsnArgGlySerPro 490
|||||
1632 CCCCAAAACGACGAGATATCATGATTTCTCTCTGTAATCGTGGAGCTCCA 1681
491 LysIleAlaSerHisGInPheSerProValAlaGlyValHisSerProke 507
|||||
1682 MAGATAGCTTCATCATCATGTTTCTCTGTTGCAAGGTGCTACTCTCCAT 1731
507 tAlaSerSerGlyAsnThrGlyAsnHisSerPheSerSerSerSerLeus 524
|||||
1732 GGCATCTTCTGGCAATACTGGGACACAGACCTTTCCACAGCTCTCTCA 1781
524 eTAlaLeuGInAlaIleSerGlyGlyValGlyThrSerLeuLeuSerThr 540
|||||
1782 GTGCCCTTGCAACCATCATGTAAGGTGTGGGACCTCCCTTTATCTACT 1831
541 LeuSerSerProGlyProLysLeuAspAsnSerProAsnMetAsnIleTh 557
|||||
1832 CTGTATCTCCAGGCCCAAAATTGATTACTCTCCCAATATGATTAATTC 1881
557 tGInProSerLysValSerAsnGInAspSerLysSerProLeuGlyPheT 574
|||||
1882 CCAACCAAGTAAAGTAAGCAATCAGAGATCCAAAGAGTCCCTGGGGCTTT 1931
574 yTcYAspGInAsnProValGlySerSerMetCysGInSerAsnSerArg 590
|||||
1932 ATTTGGACCAAAATCCAGTGAAGAGTTCAATGTCTCAGTCAAAATAGCAGA 1981
591 AspHisLeuSerAspLysGlySerLysGlySerSerValGlyGlyValaG 607
|||||
1982 GATCACCTTAGTGACAAAGAAAGTAAAGAGACAGTGTGAGGGGCGACA 2031
607 uAsnGInArgGlyProLeuGlySerLysGlyHisLysLysLeuLeuGIn 624
|||||
2032 GAATCAAAAGGGTCCCTTTGGAAGCAAAAGTCAATAAAATTACTGCAC 2081
624 euLeuThcYsSerSerAspAspArgGlyHisSerSerLeuThrAsnSer 640
|||||
2082 TACTTACCTGTTCTCTGATGACCGGGGTCATTCCTTGACCAACTCC 2131
641 ProLeuAspSerSerCysLysGlySerSerValSerValThrSerProse 657
|||||
2132 CCCCTAGATTCAAGTTGTAAGAAATCTCTGTAGTGCACAGCCCCCTC 2181
657 tGlyValSerSerSerThrSerGlyValSerSerThrSerAsnMetH 674
|||||
2182 TGGAGTCTCTCTCTCATCATCTGAGAGAGTATCCTCATCATCAATATGC 2231
674 lsgLysLeuLeuGInGlyLysHisArgIleLeuHisLysLeuLeuGIn 690
|||||
2232 ATGGGCTACTGTTTACAAAGAGACACCGGATTTTGCACAAGTTTGCAGC 2281
691 AsnGlyAsnSerProAlaGlyValAlaLysIleThrAlaGInAlaThrG 707
|||||
2282 AATGGAAATTCACAGCTGAGGTAGCCAAAGATTACTGCAAGAACCCACTG 2331
707 yLysAspThrSerSerIleThrSerCysGlyAspGlyAsnValValLysG 724
|||||

```

```

2332 GAAAGACACCAGCATATAACTTCTTGCGGACGGAATGTTGTCAACG 2381
724 InGInGInLeuSerProLysLysLysGlyuAsnAsnAlaLeuLeuArgTyr 740
|||||
2382 AGGAGCAGCTAACTCTTAAGAAAGAGAAATATATGCACTTCTTAATATAC 2431
741 LeuLeuAspArgAspAspProSerAspAlaLeuSerLysGlyLeuLeuInPr 757
|||||
2432 CTCTGACACAGGATATCTCTAGTATGCACCTCTTAAGAAACTACAGCC 2481
757 oGInValGlyGlyValAspAsnLysMetSerGInCysrThrSerSerThrI 774
|||||
2482 CCAAGTGAAGAGAGTGGATTAATTAATGATGATGATGATGATGATGATGAT 2531
774 lAProSerSerSerGInGlyLysAspProLysIleLysThrGlyThrSer 790
|||||
2532 TTCTTAAGCTCAAGTCAAGAAAGACCTTAATTAAGACAGACGACAACT 2581
791 GluGlyGlySerGlyAspLeuAspAsnLeuAspAlaIleLeuGlyLysPle 807
|||||
2582 GAAAGAGGATCTTGAGACTTGGATTAATCTAGATGCTATTTCTTGATATCT 2631
807 uThrSerSerAspPheTyrAsnAsnSerIleSerSerAsnGlySerHisL 824
|||||
2632 GACTACTTCTGACTTTTACATTAATTCATATCCTCAATGGAATGATCATC 2681
824 euGlyThrLysGInGInValPheGInGlyThrAsnSerLeuGlyLeuLys 840
|||||
2682 TGGGACTTAAGCAACAGGTGTTTCAAGGAACTAATCTCTGGGTTTGAAA 2731
841 SerSerGInSerValGInSerIleArgProProTyrAsnArgAlaValSe 857
|||||
2732 AGTTACAGTCTGTGAGTCTATCTCTCTCCATATATACCGAGCACTGTG 2781
857 rLeuAspSerProValSerValGlySerSerProProValLysAsnIleS 874
|||||
2782 TCTGATATCCCTGTTTCTGTTGGCTCAAGTCTCCAGTAATAAAATATCA 2831
874 eTAlaPheProMetLeuProLysGInProMetLeuGlyGlyLysProArg 890
|||||
2832 GTGCTTTCCCATGTTTACCAAAAGCAACCATGTTGGGTGGGATCCAAAGA 2881
891 MetMetAspSerGInGInuAsnTyrGlySerSerMetGlyGlyProAsnAr 907
|||||
2882 ATGATGATATGTCGAGAAATATATGCTCAAGTATG..... 2917
907 gAsnValThrValThrGInThrProSerSerGlyAspTyrGlyLeuProA 924
|||||
2918 .....GAGAGCTGGGGCTTTCCAA 2936
924 snSerLysAlaGlyArgMetGluProMetAsnSerAsnSerMetGlyArg 940
|||||
2937 ACTCAAAAGCCCGCAGAAATGAACTATGAACTCAATCAACCTCATGGGANA 2986
941 ProGlyGlyAspPheTyrAsnThrSerLeuProArgProAlaLeuGlyLys 957
|||||
2987 CCGAGGAGATTAATATATCTTCTTACCCAGACGCTGCACTGGGTGCTC 3036
957 rIleProThrLeuProLeuArgSerAsnSerIleProGlyAlaArgProV 974
|||||
3037 TATTTCCACATTTGCCCTTCGGGTCTAATATGATACAGAGTGGAGACGAG 3086
974 aLeuGInGInGInGInGInMetLeuGInMetArgProGlyGlyLysPro 990
|||||
3087 TATTTGCAACAGCAGCAGAGATGCTTCAATGAGGCTGCTGTAAGAACCCC 3136
991 MetGlyMetGlyAlaAsnProTyrGlyGInAlaAlaLysSerAsnGInLe 1007
|||||
3137 ATGGGAATGGGGCTATATCCCTATGTGCCAAGCAGCAGCTATACCAACT 3186
1007 uGlySerThrProAspGlyMetLeuSerMetGlyGInValSerHisGlyT 1024
|||||
3187 GGGTCTCTGGCCGATGAGCATGTTGCCATGGAACAAATTCTCATGGCA 3236

```

1024 hrglnsnaArgProleuLeuArgAsnSerLeuAspLeuValGlyPro 1040  
|||||  
3237 CTCAAATAGGCGCTCTTAGGAATCCCGATGATCTGTGGGCCA 3286  
1041 ProSerAsnLeuGlnGlnSerAspGluArgAlaLeuAspGlnLe 1057  
|||||  
3287 CCTCCAACTGGAAAGCCAGAGTGAAGAAAGACATATTGGCCAGCT 3336  
1057 uHisThrLeuSerAsnThrAspAlaThrGlyLeuGluGluLeuAsp 1074  
|||||  
3337 GCACACTCTTCTCAGCAACACAGATGCCACAGGCTGGAGAAATTGACA 3386  
1074 rgaAlaLeuGlyLeuProGluLeuValAsnGlnGlnAlaLeuGluPro 1090  
|||||  
3387 GAGCTTGGGCATTCCTGAAGCTGTCAATCAGGACAGGCAATGAGCCC 3436  
1091 LysGlnAspAlaPheGlnGlnGlnGluAlaValMetAspGlnLy 1107  
|||||  
3437 AAACAGGATGCTTCCAAAGGCCAAGAACAGCACTAATGATGATCAGAA 3486  
1107 sAlaGlyLeuTyrGlyGlnThrTyrProAlaGlnGlyProPrometGln 1124  
|||||  
3487 GGCAGGATTTATGAGACAGACATACCAGCAGGGGCTCCAAATGCAG 3536  
1124 LysGlyPheHisLeuGlnGlyGlnSerProSerPheAsnSerMetLeuAsn 1140  
|||||  
3537 GAGGCTTTCATCTTCAGGACAAATCACCATCTTTAACTCATGATGAAT 3586  
1141 GlnMetAsnGlnGlnGlnAsnPheProLeuGlnGlyMetHisProArgAl 1157  
|||||  
3587 CAGATTGAACCGAAGCAATTTCTCTCCAAAGAAATGCACCCAGCAG 3636  
1157 aAsnIleMetArgProArgThrAsnThrProLysGlnLeuArgMetGlnL 1174  
|||||  
3637 CAACATCATGAGACCCCGGACAAACACCCCAACCACTTAGAATGCAGC 3686  
1174 euGlnGlnArgLeuGlnGlnGlnPheLeuAsnGlnSerArgGlnAla 1190  
|||||  
3687 TTCAGCAGAGGCTCAGGGCCAGCAGTTTGAATCAGAGCCGACAGGCA 3736  
1191 LeuGlnLeuLysMetGluAsnProThrAlaGlyLysAlaAlaValMetArg 1207  
|||||  
3737 CTTGAAATTGAATAATGAAACCTACTAGCTGTGTGCTGCGGTGATGAG 3786  
1207 gPrometMetGlnProGlnGlnGlyPheLeuAsnAlaGlnMetValAlaG 1224  
|||||  
3787 GCCATATGATGAGCCCGCAGCGGTTTCTTAATGCTCAATGTCGCCCC 3836  
1224 LnaArgSerArgGluLeuLeuSerHisHisPheArgGlnGlnArgValAla 1240  
|||||  
3837 AACCGACAGAGAGCTGCTAAGTCATCACTCCGACAAACAGAGGTGGCT 3886  
1241 MetMetMet.....GlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1253  
|||||  
3887 ATGATGTATGACGACGACGACGACGACGACGACGACGACGACGACGAC 3936  
1253 nglnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1270  
|||||  
3937 GCAGCAGACAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3986  
1270 LaphSerProProProAsnValThrAlaSerProSerMetAspGlyLeu 1286  
|||||  
3987 CCTTCACGCCCACTCTTAATGTGACTGCTTCCCGACATGATGAGGCTT 4036  
1287 LeuAlaGlyProThrMetProGlnAlaProProGlnGlnPheProTyrGly 1303  
|||||  
4037 TTGGCAGGACCCCAATGCCACAGCTCTCCGACACAGTTTCCATATCA 4086  
1303 nProAsnTyrGlyMetGlyGlnGlnProAspProAlaPheGlyArgValS 1320  
|||||  
4087 ACCAAATTATGGAATGGACACAACAGATCCAGCTTTGGTCGAGTGT 4136

1320 erSerProProAsnAlaMetMetSerSerArgMetGlyProSerGlnAsn 1336  
|||||  
4137 CTAGTCTCTCCCAATGCAATGATGTGCTCAAGAAATGGCTCCCTCCAGAAAT 4186  
1337 PrometMetGlnHisProGlnAlaAlaSerTleTyrGlnSerSerGluMe 1353  
|||||  
4187 CCCATGATGCAACACCCGCAAGGCTGCATCCATCTATAGTCCTCAGAAAT 4236  
1353 tLysGlyTyrProSerGlyAsnLeuAlaArgAsnSerSerPheSerGln 1370  
|||||  
4237 GAAGGCTGGGCATCAGGAATTTGGCCAGAGAACAGCTCTTTTCCAGC 4286  
1370 lnglnPheAlaHisGlnGlyAsnProAlaValTyrSerMetValHisMet 1386  
|||||  
4287 AGCAGTTTGGCCACAGGAGGAATCTGACGTATAGTATGATGTCACATG 4336  
1387 AsnGlySerSerGlyHisMetGlyGlnMetAsnMetAsnPrometProm 1403  
|||||  
4337 AATGGCAGCAAGTGTCAATGGAGACAGATGAACATGAAACCCCATGCCAT 4386  
1403 tSerGlyMetProMetGlyProAspGlnLysTyrCys 1415  
|||||  
4387 GTCGGCATGGCTATGGGTCTCTGATCAGAAATACTGC 4423



**This Page Blank (uspto)**



151 AspleValAsnThrSerValTyrAsnIleuHisGluAspArgLy 167  
 |||||  
 635 GACCTGGTTACACAAAGTGTACATATCTTACATGAAAGAACGAAA 684  
 167 sAspPheLeuLysAsnLeuProLysSerThrValAsnGlyValSerTrp 184  
 |||||  
 685 GGATTTTCTTAAGAAATTTACCAAAATCTACAGTTAAAGAGTTTCTGGA 734  
 184 hAsnGluProGlnArgGlnLysSerHisThrPheAsnGlySerMetLeu 200  
 |||||  
 735 CAATGAGAGCCCAAGACAAAAAGCCATACATTTAATTGGCGTAGTTGG 784  
 201 MetLysThrProHisAspIleLeuGluAspIleAsnAlaSerProGlu 217  
 |||||  
 785 ATGAAGAACCAACATGATATCTGGAAGACATAAACGCGAGTCTGAAT 834  
 217 ValGlnArgTyrGluThrMetGlnCysPheAlaLeuSerGlnProArg 234  
 |||||  
 835 GGCAGAGATATGAACCAATGCGAGTGTGGCTGTCTCAGCCAGAG 884  
 234 lAspMetGluGluGlnGluAspLeuGlnSerCysMetIleCysValAla 250  
 |||||  
 885 CATTGATGAGAGAGGAGGAAGATTTCATCTTGTATGATCTGTGCGCA 934  
 251 ArgArgIleThrThrGlyGluArgThrPheProSerAsnProGluSerPh 267  
 |||||  
 935 CCGCGCATTAACACAGAGAGAAAGAACATTTCCATCAACCCGAGAGCT 984  
 267 eIleThrArgHisAspLeuSerGlyLysValValAsnIleAspThrAsn 284  
 |||||  
 985 TATTACCAACATGATCTTTCAGAAAGGTGTCAATATAGATACAAAT 1034  
 284 eLeuArgSerSerMetArgProGlyPheGluAspIleIleArgArgCys 300  
 |||||  
 1035 CACTGAGATCCCTCCATGAGCGCTTGTGAAGATTAATCCGAAAGTGT 1084  
 301 lIleGlnArgPhePheSerLeuAsnAspGlyGlnSerTrpSerGlnLysAr 317  
 |||||  
 1085 ATTCAGAGATTTTGTAGCTAAATGATGGGAGTGTGCTCCAGAGAACG 1134  
 317 gHisTyrGlnGluAlaTyrLeuAsnGlnLysAlaGluThrProValTyrA 334  
 |||||  
 1135 TCATCTACAGAAAGCTTATCTTATATGCGCATGACAGAACCCCGATATATC 1184  
 334 rPheSerLeuAlaAspGlyThrIleValThrAlaGlnThrLysSerLys 350  
 |||||  
 1185 GATTTCTGTTGGCTGATGGAACTATAGTGCCTGCACAGACAAAAGCATA 1234  
 351 LeuPheArgAsnProValThrAsnAspArgHisGlyPheValSerThrH 367  
 |||||  
 1235 CTCTTCGGAATCTGTAACAATGATGACATGGCTTGTCTCAACCCA 1284  
 367 sPheLeuGlnArgGluGlnAsnGlyTyrArgProAsnProAsnProValG 384  
 |||||  
 1285 CTTCTTCAGAGAGACAGAAATGATATAGACCAAAACCAATCTCTGTG 1334  
 384 lGlnGlyIleArgProPrometAlaGlyCysAsnSerSerValGlyGly 400  
 |||||  
 1335 GACAAAGGATAGACCACTATGGGTGATGCAACAGTTGCGTAGGCGCG 1384  
 401 MetSerMetSerProAsnGlnGlyLeuGlnMetProSerSerArgAlaI 417  
 |||||  
 1385 ATGAGATATCTGCGCAAAACCAAGCTTACAGATGCCAGAGCGAGGCCCTA 1434  
 417 rGlyLeuAlaAspProSerThrThrGlyGlnMetSerGlyAlaArgTyrG 434  
 |||||  
 1435 TGGCTTGGAGACCTTAGCAACCAAGGAGATGATGAGCTAGGTATAG 1484  
 434 lGlyLysSerAsnIleAlaSerLeuThrProGlyProGlyMetGlnSer 450  
 |||||  
 1485 GGGGTTCAGATACATAGCTTATGATGACCCGTGGCCAGGACGACATCA 1534

451 ProSerSerTyrGlnAsnAsnAsnTyrGlyLeuAsnMetSerSerPro 467  
 |||||  
 1535 CCATCTTCTTACCAGAACACAACTATGGCTCAATAGTAGCCCCC 1584  
 467 oHisGlySerProGlyLeuAlaProAsnGlnGlnAsnIleMetIleSerP 484  
 |||||  
 1585 ACATGGGAGTCTCTGCTTGGCCCCCAACGACGAGAAATATCATGATTTTC 1634  
 484 rArgAsnArgGlySerProLysIleAlaSerHisGlnPheSerProVal 500  
 |||||  
 1635 CTGATATGCTGGAGAGTCCAAAGATAGCTCACAATCAGATTTCTCTGTT 1684  
 501 AlaGlyValHisSerPrometAlaSerSerGlyAsnThrGlyAsnHis 517  
 |||||  
 1685 GCAGGTGCACCTCTCCATGGCATCTTGTGCAATACGGGAGAACACAG 1734  
 517 rPheSerSerSerLeuSerAlaLeuGlnAlaIleSerGlnGlyValG 534  
 |||||  
 1735 CTTTTCACAGAGCTCTCTCAGTGCCCTGCAGGCAATCAGTAAGGTGTGG 1784  
 534 lYThrSerLeuLeuSerThrLeuSerSerProGlyProLysLeuAspAsn 550  
 |||||  
 1785 GGACTTCCCTTTTATCTACTCTGTCAATCAGAGGCCCAATTTGGATATC 1834  
 551 SerProAsnMetAsnIleThrGlnProSerLysValSerAsnGlnAsp 567  
 |||||  
 1835 TCTCCCAATATGAAATATTAACCAACCAAGTAAAGTAAGCAATCAGATTC 1884  
 567 rLysSerProLeuGlyPheTyrCysAspGlnAsnProValGluSerSer 584  
 |||||  
 1885 CAAGAGTCCCTCTGGGCTTTTATTGGCACCMAAATCCAGTGGAGAGTTCAA 1934  
 584 eCysGlnSerAsnSerArgAspHisLeuSerAspLysGluSerLysGlu 600  
 |||||  
 1935 TGTGTACGTAATATAGCAGAGATCACCTCAGTGACAAAGAAAGTAAGAG 1984  
 601 SerSerValGluGlyAlaGluAsnGlnArgGlyProLeuGluSerLysG 617  
 |||||  
 1985 AGCAGGTGTAGAGGGGACAGAAATCAAGGGGCTCTTGGAAAGCAAAAG 2034  
 617 yHisLysLysLeuLeuGlnLeuLeuThrCysSerSerAspAspArgGlyH 634  
 |||||  
 2035 TCATTAATAATTAATCTGACATTAATCTTCTTGTATGACGGGGGCTC 2084  
 634 lAspSerLeuThrAsnSerProLeuAspSerSerCysLysGluSerSer 650  
 |||||  
 2085 ATTCCTCTTGACCAACATCCCTCGATTAATCAAGTGTAAAGAAATCTTCT 2134  
 651 ValSerValThrSerProSerGlyValSerSerSerThrSerGlyVal 667  
 |||||  
 2135 GTTAGTGTACACAGCCCCCTGGAGTCTCTCTCTTACATCTGGAGGAGT 2184  
 667 lSerSerThrSerAsnMetHisGlySerLeuLeuGlnGlyLysHisArgI 684  
 |||||  
 2185 ATCTCTTACATCAATATGATGATGGTCACTGTTTACAAAGAAAGCACCGGA 2234  
 684 lLeuHisLysLeuLeuGlnAsnGlyAsnSerProAlaGluValAlaLys 700  
 |||||  
 2235 TTTTGCACAAAGTTGTCGAGAAATGGGAATTCACCACTGAGTAGGCAAG 2284  
 701 lIleThrAlaGlnAlaThrGlyLysAspThrSerSerIleThrSerCysG 717  
 |||||  
 2285 ATTAATGCAAGAACCTCTGGGAAAGACACACAGATATTAATCTTGTGG 2334  
 717 yAspGlyAsnValValLysGlnGlnLeuLeuSerProLysLysGlyAla 734  
 |||||  
 2335 GGACGGAATTTGTCTCAACAGAGAGCAGCTAAGTCTTAAGAAAGAGGAGA 2384  
 734 snAsnAlaLeuLeuArgTyrLeuLeuAspArgAspAspProSerAspAla 750  
 |||||  
 2385 ATATATCAGACTTCTTAGATACCTGTCGACAGGAGATGATCTTGTATGCA 2434  
 751 LeuSerLysGluLeuGlnProGlnValGluGlyValAspAsnLysMetSe 767

```

|||||
2435 CTCTCTAAGACAGACAGCCCAAGTGGAGAGATGATATAATAAGAG 2484
767 rGlnCysThrSerSerThrIleProSerSerSerGlnGluLysAspProL 784
|||||
2485 TCAGTGCACAGCCTCCACCATCTCTACCTCAAGTCAAGACAAAGACCCTA 2534
784 ysiIeUysThrGluThrSerGlnGluGlySerGlyAspLysAspAsnLeu 800
|||||
2535 AAATTAAAGACAGACAGATGAGAGGATCTGGAGACTTGGATTAATCTA 2584
801 AspaIaIleUeuGlyAspLeuThrSerSerAspPheTyrAsnAsnSerI 817
|||||
2585 GATGCTATTCTTGCTGATCTGACACTGACTTGACTTTACATAATTCAT 2634
817 eSerSerAsnGlySerHisLeuGlyThrLysGlnGlnValPheGlnGlyT 834
|||||
2635 ATCCTCAAAATGCTAGTCACTGGGAGCTAAAGCAACAGGTGTTCAAGAA 2684
834 hirsAsnSerLeuGlyLeuLysSerSerGlnSerValGlnSerIleArgPro 850
|||||
2685 CTAAATCTCTGGGTTTGAAGATTCACAGTCTGTGACGTCTATTGCTCT 2734
851 ProTyrAsnArgAlaValSerLeuAspSerProValSerValGlySerSe 867
|||||
2735 CCATATAACGAGCAGCTGCTCTGTGATAGCCCTTTCTGTGGCTCAG 2784
867 rProValLysAsnIleSerAlaPheProMetLeuProLysGlnProM 884
|||||
2785 TCCCTCAGTAAATAATACGTGCTTCCCATGTATCCAAAGCAACCCA 2834
884 etLeuGlyGlyAsnProArgMetMetAspSerGlnGluAsnTyrGlySer 900
|||||
2835 TGTGGGTGGGAATCCAGATGATGATAGTACGAAATTAATGCTCA 2884
901 SerMetGlyGlyProAsnArgAsnValThrValThrGlnThrProSerSe 917
|||||
2885 AGTATGGGTGGGCCAAACCCAAATGTGACTGTGCTCAGACTCCTTCTC 2934
917 rGlyAspTyrPglLysLeuProAsnSerLysAlaGlyArgMetGluProMetA 934
|||||
2935 AGGAGACTGGGGCTTACCAAACTCAAAAGCCGCGCAGATGGAACCTTA 2984
934 snSerAsnSerMetGlyArgProGlyGlyAspTyrAsnThrSerLeuPro 950
|||||
2985 ATTCAAACTCCATGGGAAGACAGACAGAGAGATTAATTAATCTTTTACC 3034
951 ArgProAlaLeuGlyGlySerIleProThrLeuProLeuArgSerAsnSe 967
|||||
3035 AGACCTGCACCTGGGTGGCTCTATTCCACAGTGGCTCTTCGGCTTAATG 3084
967 rIleProGlyAlaArgProValLeuGlnGlnGlnGlnMetLeuGlnM 984
|||||
3085 CATACCAAGGTGCGAGACAGATTTGCAACAGACAGACAGATGCTTCAAA 3134
984 etArgProGlyGluIleProMetGlyMetGlyAlaAsnProTyrGlyGln 1000
|||||
3135 TGAGGCTGTGGAAATCCCATGGGAATGGGGGCTTAATCCCTATGGCCAA 3184
1001 AlaIaIaIaSerAsnGlnLeuGlySerThrProAspGlyMetLeuSerMe 1017
|||||
3185 GCACGACGATCTTAACCAACTGGGTTCCTGGGCCCATGGCATGTGTCCAT 3234
1017 tGluGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerL 1034
|||||
3235 GGAACAAGTTTCTCATGCGACTCAAAATAGGCCCTTTCTTAGAATGCC 3284
1034 euAspAspLeuValGlyProProSerAsnLeuGlyGlnGlnSerAspGlu 1050
|||||
3285 TGGATGTACTTGTGGGCGACCTTCCAAACTGGAAGGCGACAGAGACGAA 3334
1051 ArgAlaIaIaSerAsnGlnLeuHisThrLeuLeuSerAsnThrAspAlaTh 1067
|||||
3335 AGACATTATTGGACGACGCTGACACTCTTCTCAGCAACACAGATGGCAC 3384
1067 rGlyLeuGlnGluIleAspArgAlaLeuGlyIleProGluLeuValAsnG 1084
|||||
3385 AGGCTGTGAAGAAATTCAGACAGCTTGGCATTCGAACTGTCTCATC 3434
1084 lngGlyAlaIaLeuGluProLysGlnAspAlaPheGlnGlnGlnAlaI 1100
|||||
3435 AGGACACAGCATTTAGACCCCAACAGCATGCTTTCCAAAGCCCAAGAA 3484
1101 AlaValMetMetAspGlnLysAlaGlyLeuTyrGlyGlnThrTyrProAl 1117
|||||
3485 GCAGTAAATGATGATCAGAAAGGAGATTTATGACAGACATACCAGC 3534
1117 aGlnGlyProProMetGlnGlyGlyPheHisLeuGlnGlyGlnSerProS 1134
|||||
3535 ACAGGGGCTTCCATTCAGAGAGGCTTTCATCTTCAGGGCAATACCAT 3584
1134 erPheAsnSerMetMetAsnGlnMetAsnGlnGlnGlnLysAsnPheProLeu 1150
|||||
3585 CTTTAATCTATGATGATATCAGATGAACGACAGGCAATTTTCTCTC 3634
1151 GlnGlyMetHisProArgAlaAsnIleMetArgProArgThrAsnThrPr 1167
|||||
3635 CAAGGAATGCAACCCAGACGCCAATCATGAGACCCCGGCAAAACACCC 3684
1167 oLysGlnLeuArgMetGlnLeuGlnGlnArgLeuGlnGlyGlnGlnPheL 1184
|||||
3685 CAACCAACTTGAATGACAGCTTCAGCAGAGGCTCAGAGGCCAGCAGTTT 3734
1184 euAsnGlnSerArgGlnAlaLeuGlnLeuLysMetGluAsnProThrAla 1200
|||||
3735 TGAATCAGAGCGCAGAGGCACTTAATTAATGAATGAATGAATCCATGCT 3784
1201 GlyGlyAlaAlaValMetArgProMetMetGlnProGln.....G 1214
|||||
3785 GGTGTGCTGCGGTGATGAGGCTATGATGAGGCCAGCCAGGTGAGCTCCA 3834
1214 nGlyPheLeuAsnAlaGlnMetValAlaGlnArgSerArgGluLeuLeuS 1231
|||||
3835 GGGTTTTCTTAATGCTCAAAATGGTGCGCCCAACGACGAGAGAGCTGTA 3884
1231 erHisHisPheArgGlnArgGlnArgValAlaMetMetMet..... 1243
|||||
3885 GTCATCACTTCCGACACAGAGGGTGGCTATGATGATGACAGACAGCAA 3934
1244 ...GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1259
|||||
3935 CAGCAGCAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 3984
1259 nGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1276
|||||
3985 ACAGCAACAGCAGCAGCAGCAGCAACCCAGGCTTCAGGCCACCTCCTA 4034
1276 snValThrAlaSerProSerMetAspGlyLeuLeuAlaGlyProThrMet 1292
|||||
4035 ATGTGACGTGTTCCCGCAGCATGATGGGCTTTTGGAGAGGCCACCATG 4084
1293 ProGlnAlaProProGlnGlnPheProTyrGlnProAsnTyrGlyMetG 1309
|||||
4085 CCACAGCTCTCCGCAACAGATTTCATATCAACCAATTTATGGAATGG 4134
1309 yGlnGlnProAspProAlaPheGlyArgValSerSerProProAsnAlaM 1336
|||||
4135 ACAACAACCAATCAGAGCTTTGGTCGAGTGTATGTCCTCCCAATGAA 4184
1336 etMetSerSerArgMetGlyProSerGlnAsnProMetMetGlnHisPro 1342
|||||
4185 TGATGTGTCTCAAGATGGTCCCTCCAGAAATCCCATGATGCCAACACCC 4234
1343 GlnAlaIaIaSerIleTyrGlnSerSerGlnMetLysGlyTyrProSerG 1359
|||||
4235 CAGGCTGCATCATCTATCAGTCTCCTCAGAAATGAAGGCGTGGCATCAG 4284

```

1359 yAsnLeuAlaArgAsnSerSerPheSerGlnGlnInpHeAlaHisGln 1376  
1376 yAsnProAlaValTyrSerMetValHisMetAsnGlySerSerGlyHis 1392  
4285 AATTTGGCCAGGACGCTCTTTCCGACGAGTTGGCCACAGG 4334  
1376 yAsnProAlaValTyrSerMetValHisMetAsnGlySerSerGlyHis 1392  
4335 GGATCTCTCAGCTGTAGTATGTGTGACATGTAATGGACAGTGGTAC 4384  
1393 MetGlyMetAsnMetAsnPrometPrometSerGlyMetPrometG 1409  
4385 ATGGGACAGATGAACATGAACCCCATGCCCATGTGTGGCATGCCATTGG 4434  
1409 yProAspGlnLysTyrCys 1415  
4435 TCCTGATCAGAAATACTGC 4453

seq\_name: /cgn2\_6/ptodata/2/pna/us09\_NEW\_COMB.seq:us-09-196-296B-3

seq\_documentation\_block:

Sequence 3, Application US/09196296B  
GENERAL INFORMATION:  
APPLICANT: Suen, Chen-Shian  
APPLICANT: Fraill, Donald E.  
APPLICANT: Lyttle, Richard C.  
TITLE OF INVENTION: Cloning and Expression of a Nuclear  
RECEPTOR COACTIVATOR PROTEINS AND USES THEREOF  
FILE REFERENCE: 0650/01376  
CURRENT APPLICATION NUMBER: US/09/196,296B  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 3  
LENGTH: 3361  
TYPE: DNA  
ORGANISM: Murine  
US-09-196-296B-3

Alignment\_scores:  
Quality: 480.00 Length: 732  
Ratio: 2.840 Gaps: 4  
Percent Similarity: 23.087 Percent Identity: 21.585

Alignment\_block:

US-09-041-994-2 x US-09-196-296B-3 ..

Align seg 1/1 to: US-09-196-296B-3 from: 1 to: 3361

7 AsnLeuAspProLeuAlaSerAspSerArgLysArgLysLeuProCys 23  
4 TCTTTGGATCCGCTGGCCGCTGTGATCTCGGAACGCAAACTGCCCTGTGA 53  
23 pThrProGlyGlnGly..... 28  
54 TCTCTCAGACAGAGGGGTAGTATCATCTCCATTTCAGATGTTTCTCA 103  
28 ..... 28  
104 GGCCAGGACTAATCAAAAGTGACACTGAACAAGAAATTTCTTCACTGA 153  
28 ..... 28  
154 TCCTTTGTGCCACACCTTGATGCTTCTTGTGGCTAAGTTAGTTAAT 203  
28 ..... 28  
204 TCACACAAACTAATGCATATATGGGGGGAATTCGCTTAAGTTTCT 253  
29 .....LeuThrCysSer 32  
254 TCAGAGGTAGTGTGTGCTTTTCTCCGCTTCCAGGCTTGCTACAGT 303

33 GlyLysArgArgLysGlnGlnLysSerLysTyrIleGlnLysLeuAl 49  
304 GTGAGAGAGTGGCGGAGGAGGAGAGAGAGTATGATGAGAGCTGGC 353  
49 AsnLeuIleSerAlaAsnLeuSerAspIleAspAsnProAlaValLysP 66  
354 AGAGCTCATCTCTGCAAAATCTCAGGCATCGCAACTTCAATGTCAAG 403  
66 roAspLysCysAlaIleLeuLysGluThrValArgGlnIleArgGlnIle 82  
404 CAGATTAATGTGCTCATCTTAAGAGACAGTGAACAGATAGCGCAATA 453  
83 LysGlnGlnGly..... 86  
454 AAAGAACAAAGGTAAACACAGAGTCAGAAAAAACCCAAAGTTGGACTGT 503  
86 ..... 86  
504 AATATGTATGAATCCTCCCTCCCTCCCTTCTGTGATCAGAGACTTAATA 553  
86 ..... 86  
554 GAGCTCAGCTGTGTAAAGTAAGTACTTACCAAGTCCATGAGTTGACCA 603  
86 ..... 86  
604 GCTCTTACTTTAAGAACAATTACAGAAATGGCAGAAAGCCAGATGTC 653  
86 ..... 86  
654 TTCAGAACCGGAGATCTGTGTGGAGCAAGCTAAGGCCCTTAAGCTAC 703  
86 ..... 86  
704 ATGGAGCAGAGAGGAGGAGCCAGGAGTCTTCTTACTTCTGGAAC 753  
86 ..... 86  
754 CTGAGGCTTCTTATCAGTAGTGTGCTTGTGATATATGTACA 803  
86 ..... 86  
804 TACTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 853  
86 ..... 86  
854 CTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 903  
86 ..... 86  
904 TGTGTAAAGGCTGAGGTAGTGTCTCAACAGACGCTTCTGTGTGT 953  
87 .....Lys,ThrIleSerAsnAspAspValGlnLysAlaAspVal 100  
954 CTTACAGAAAACTATTTCCAGATGATGATGTTCAAAAAGCTATGTG 1003  
101 SerSerThrGlyGlnGlyValIleAspLysAspSerLeuGlyProLeu 117  
1004 TCTTTCACAGGCGAGGAGTCAATTGATTAAGACTCTTGAAGCCGCTTT 1053  
117 uLeuGln..... 119  
1054 ACTACAGGTAGGTGTGAGGATTTGTGGCTGTCCAAAGCATGATGCTTTG 1103  
119 ..... 119  
1104 GTTGTGTACTCTGTGATTTATGTGTGTGGGCGAGGAGATTGGAAT 1153  
119 ..... 119  
1154 GAGAGATAGTTTCTACCCCGTGAGAGAGACTCAGTATTAAGACATC 1203  
119 ..... 119



```

2104 GTACACAGAGAGACCTGGTTAAACAAAGTGTCACAGACATCTACATG 2153
163 luguasparglysasphenleuylasnpProlyserThr 177
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2154 AGCAAGACCGGAAGATTTTCTTAAACACTTACCAAAATCCACA 2197
seq_name: /cgn2_6/plodata/2/pna/PCI_NEW_COMB.seq:PCI-US01-01339-9844
seq_documentation_block:
; Sequence 9844, Application PC/TUS0101339
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01339
; NUMBER OF SEQ ID NOS: 2001-03-17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 9844
; LENGTH: 2487
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US01-01339-9844

alignment_scores:
Quality: 299.50 Length: 1196
Ratio: 0.543 Gaps: 54
Percent Similarity: 46.154 Percent Identity: 21.572

alignment_block:
US-09-041-994-2 x PCT-US01-01339-9844/rev ..
Align seg 1/1 to reverse of: PCT-US01-01339-9844 from: 1 to: 2487

327 HislaqluthrProValThrArg.....PheSerleuAl 338
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
14073 CACCTGTGAACACCTCCAAAGAGAGAGAGATCTGGCTCATCTTCAGAGCG 14024
338 aaspglythrIleValThrIaqluthrIlysserIlyseuPheArgasp 355
14023 GAAAAACAAATCCAGAACATCTCAAGAAAGAGAGAGGTCATTCACCC 13974
355 ro.....ValThr 357
|||
13973 CAGAAATGAAGAAATCTCGCATTTCTTCAAGGGGAGGAGGCTCTCTCT 13924
358 AsnAsparghIsglyPheValSerThrHisPheleuGlnArgGluGlnAs 374
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13923 TCACCAAGGTCACCAAGCAAAATCTCGCTTGTCTTGGAGGCGCCTTTC 13874
374 nglYThrArgProAsnProAsnProValGlyGln..... 385
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13873 AGGGTCTTCCCATGCCCTTAAGCAAAAGTCACGACACCCAGCCAGCGCA 13824
386 .....GlyIleArgProPromet 391
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13823 GTGCGTCTGGATCCTCCCAAACTAAAGTAAATCTAAGAACGACACCCAGA 13774
392 AlaIglysasnserserValGlyGlyMetSerMetSerProAsnGlnG 408
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13773 CGCAGTGTCCGTCAGTTCT.....TCGCCGCCACCTAAACAGAA 13736
408 yLeuGlnMetProserSerArgAlaThrGlyLeuAlaAspPro..... 422
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13735 ATCTAAGACACCATCAAGACAAAGTATTCACCTGATATCTCTCATCTTA 13686
423 .....SerThrThrglyGlnMetSerGlyAlaIaThrGlyIlyser 436
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13685 AAGTGAATATCTGGAACACCAAGGAGGAGGTCATTAACAGTCCCGAG 13636
437 SerAsnIleAlaSerIleThr..... 443
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

714	reserCysglYaspglyasrValValLysInclugInleu.....	727
715	.....	
716	.....	
717	.....	
718	.....	
719	.....	
720	.....	
721	.....	
722	.....	
723	.....	
724	.....	
725	.....	
726	.....	
727	.....	
728	.....	
729	.....	
730	.....	
731	.....	
732	.....	
733	.....	
734	.....	
735	.....	
736	.....	
737	.....	
738	.....	
739	.....	
740	.....	
741	.....	
742	.....	
743	.....	
744	.....	
745	.....	
746	.....	
747	.....	
748	.....	
749	.....	
750	.....	
751	.....	
752	.....	
753	.....	
754	.....	
755	.....	
756	.....	
757	.....	
758	.....	
759	.....	
760	.....	
761	.....	
762	.....	
763	.....	
764	.....	
765	.....	
766	.....	
767	.....	
768	.....	
769	.....	
770	.....	
771	.....	
772	.....	
773	.....	
774	.....	
775	.....	
776	.....	
777	.....	
778	.....	
779	.....	
780	.....	
781	.....	
782	.....	
783	.....	
784	.....	
785	.....	
786	.....	
787	.....	
788	.....	
789	.....	
790	.....	
791	.....	
792	.....	
793	.....	
794	.....	
795	.....	
796	.....	
797	.....	
798	.....	
799	.....	
800	.....	
801	.....	
802	.....	
803	.....	
804	.....	
805	.....	
806	.....	
807	.....	
808	.....	
809	.....	
810	.....	
811	.....	
812	.....	
813	.....	
814	.....	
815	.....	
816	.....	
817	.....	
818	.....	
819	.....	
820	.....	
821	.....	
822	.....	
823	.....	
824	.....	
825	.....	
826	.....	
827	.....	
828	.....	
829	.....	
830	.....	
831	.....	
832	.....	
833	.....	
834	.....	
835	.....	
836	.....	
837	.....	
838	.....	
839	.....	
840	.....	
841	.....	
842	.....	
843	.....	
844	.....	
845	.....	
846	.....	
847	.....	
848	.....	
849	.....	
850	.....	
851	.....	
852	.....	
853	.....	
854	.....	
855	.....	
856	.....	
857	.....	
858	.....	

```

968 eProGIyAlaArProValLeuGlnGlnGlnGlnMetLeuGlnMet 985
      ::::: |||
11953 TTCGCAATCTTCCT..... 11938
985 rProGIyGluIleProMetGlyMetGlyAlaAsnProTyrglyGlnAla 1001
      ||::: ||::: ||:::
11937 .....GAATGAAAGATGTTTACCAGACACTCCA..... 11908
1002 AlaAlaSerAsnGlnLeuGlySerTrpPro.....AspGlyMetLe 1015
      ::::: ||::: |||
11907 TCAAGGAAACAGGCTCGCTTCCTCCAGACTTACAGATGGG.... 11863
1015 uSerMetGluGlnAlaSerHisGlyThrGlnAsnArgProLeuArg 1032
      ||::: ||::: ||:::
11862 .....TCTGGACTCTCCTCGAGGCACAGCCTGTCTG 11832
1032 snSer.....LeuAspAspLeuValGlyProProSerAsnLeuGlu 1045
      ::::: ||::: ||::: ||:::
11831 GGTCTCTCTCGAATGAATATATACCTAGACGCCATCTAGAGGAGA 11782
1046 GlyGlnSerAspGluArgAlaLeuLeuAspGlnLeuHisThrLeuSe 1062
      ::::: |||
11781 AGCGAATGTGTAT.....TC 11768
1062 rAsnThrAspAlaThrGlyLeuGluGluIle.....AspArg 1075
      ||::: ||::: ||::: ||:::
11767 TTCCCGAGAACCGAAAGCTTTGCTCCTCAGACTCTAGGCCGAGAGTCTT 11718
1075 lAlaGlyIleProGluLeuValAsnGlnGlnAlaLeuGluProlys 1091
      ::::: ||::: ||::: ||::: ||:::
11717 CTCATCATCTCCAGAGCTCAACAAC.....AGTGTCTTACCCCTCAG 11674
1092 GlnAspAlaPheGlnGlyGlnAlaAlaValMetMetAspGlnLysAl 1108
      ::::: ||::: ||::: ||::: ||:::
11673 AGAGAAAGA...ACGGGTCCAGAAATCATCAGTT.....GATCGAAAC 11633
1108 a.....GlyLeuTyrglyGlnThrTyrglyProAlaGln 1119
      ||::: ||::: ||::: ||::: ||:::
11632 TGTGGCTCGGACTCCCTGGGGCAGAGAGTCTGGGATCTCTCAAG 11583
1119 lProProMetGlnGlyGlyPheHisLeuGlnGlnGlnSerProSerPhe 1135
      ::::: ||::: ||::: ||::: ||:::
11582 AACTTGATGTGAACCCAGTGCATCCCTCAGGAAGAGATGATCAAC 11533
1136 AsnSer.....MetMetAsnGlnMetAs 1143
      ::::: |||
11532 TCTTCTCCAGATTTAAAGCCAGACAGAACCCCACTTCGCGAGAGAG 11483
1143 nGlnGlnGlyAsnPheProLeuGlnGlyMetHisProArgAlaAsnIle 1160
      ::::: ||::: |||
11482 TCGGCTGGATCATCTCCAGAGTTGACAGCAAACTTCGACTATCCCTC 11433
1160 eTarGProArgThrAsnThrProLysGlnLeuArgMetGlnLeuGln 1176
      ||::: ||::: ||::: ||::: ||:::
11432 GGGCAGTAGGTCTGTCTCTCCCTGAAAGTGAAGATAACCAAGACAC 11383
1177 ArgLeuGlnGlyGln..... 1182
      ::::: ||::: ||:::
11382 GCACCCAGGACAGAGTGTCTGATTCCTCTCGAACCTAAAGCTTC 11333
1183 .....PheLeuAsnGlnSerArgGlnAlaLeuGluLeuLysM 1195
      ::::: ||::: ||::: ||::: ||:::
11332 AGCCCCTCGGGCCCTCCAGACAGAGATCAAGTTTCATCAAGCAAG 11283
1195 eGluGlnProThrAlaGlyGlyAlaAlaValMetArgProMetMetG 1211
      ::::: ||::: ||::: ||::: ||::: ||:::
11282 GCAAGAGGCCCTTCTCTGGAAGAGACAGACTACGAGTCTCTCTCTAA 11233
1211 nProGlnGlnGlyPheLeuAsnAlaGlnMetValAla..... 1223
      ::::: |||
11332 CATCCGCCCA.....AATCCAGAACTCTCTGACAGAGTTCCAGCTC 11192
1224 .....GlnArgSerArgGluLeuLeuSerHisThrPheArgGlnGlnArg 1238

```

```

11191 ATCACCAGAGCCCAAGACCAAGTCTCTACACCACT..... 11155
      ||| ||| ||| |||||
1239 ValAlaMetMetMetGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1255
      ::::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::
11154 .....CGAGTGCAGCTCTCATCATCTCCGGA 11126
1255 nGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1272
      ::::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::
11125 GCTAACCAAGGAAGGACAGACTGTCCGTAAGAAAGCGCTCTCTCATCT 11076
1272 erPro.....ProProAsnValThrAlaSerProSer 1282
      ||||| ||||| ||||| ||||| ||||| |||||
11075 CACCAAGAACTCGCTTAGAACTCCCCAGAGCACCGGAGAAATCTCTCA 11026
1283 MetAspGlyLeuAlaGlyProThrMetProGlnAlaProGlnGln 1299
      ::::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::
11025 GTG..... 11023
1299 nPheProTyrglnProAsnTyrglyMetGlyGlnGlnProAspProAla 1315
      ::::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::
11022 .....TCTTCCCGGAGCCAGCCG 11004
1316 .....PheGlyArgValSerSerProProAsnAla 1325
11003 AAAAAATCGAGCTCTTCACGCCGAGCGCGCTCAGCTCATCTCAGCACT 10954
1326 MetMetSerSerArgMetGlyProSerGlnAsnPro 1337
      ::||| ||| ||| |||
10953 AAGACAACTCTCAGAGAGGCGGCTCTCTCTCGCCA 10918
seq.name: /cgn2.6/ptodta/2/pna/PCT_NEW_COMB.seq:PCT-US01-01339-7281
seq.documentat_block:
; Sequence 7281, Application PC/TU0101339
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01339
; CURRENT FILING DATE: 2001-03-17
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7281
; LENGTH: 13555
; TYPE: DNA
; ORGANISM: Homo sapiens
; PCT-US01-01339-7281
alignment_scores:
Quality: 226.50 Length: 1330
Ratio: 0.387 Gaps: 67
Percent Similarity: 44.060 Percent Identity: 21.053
alignment_block:
US-09-041-994-2 x PCT-US01-01339-7281 ..
Align seg 1/1 to: PCT-US01-01339-7281 from: 1 to: 13555
330 ThrProValTyArgPheSerLeuAlaAspGlyThrIleValThrAlaG 346
      ::||| ::::: ||::: ||::: ||::: ||::: ||
5108 TCCCTCAAGTCAAGAAAGATACCAACCATTTCTCCAGATTCTCCACA 5157
346 nThrLysSerLysLeu...PheArgAsnProValThrAsnAspArgHis 362
      ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::
5158 AAACCCAGAGACCTCAGCTCAAGAGGCGCTTTAGTACACCTGCTGCT 5207
362 lPheValSerThrHisPheLeu.....GlnArgGlnGlnAsnGly 375
      ::||| ||| ||| ||| ||| ||| ||| |||
5208 TATCTCTTCAACATCAGTCTCTCTGAGTACCTCTTCAAAAGACT 5257
376 TyArgProAsnProAsnPro.....ValGlyGlnGlyI 387

```

```

5258 GCGGGTCCCAACACCCCCAGATTTTCCCATTTCTCTGGGCTCTCATCT 5307
      |||||  |||  ::|||:::
387 eaIgpProMetalIaIcysAsnSerSerValIglYMetSerMetS 404
      :  |||  ::|||:::
5308 TGCACCTTTACAT 5349
      :  |||  |||||:::
404 erProAsnIglYLeuGlnMetProSerSerAraIaIaIYglYLeuAla 420
      :  |||  |||||:::
5349 TAGGTCAACAGGT 5377
      :  |||  |||||:::
421 AspPro 5377
      :  |||  |||||:::
5378 GACCCTACATGAAGACATTTCTGTAGATCATTTCTCCACAGGCGCTC 5427
      :  |||  |||||:::
432 gTyIglYglYSerSerAsnIleAlaSerLeuThrProGlyProGly 447
      :  |||  |||||:::
5428 TTTATCTTTCAGAGATCTGTATCTCTCCCTTCCCTCCAGAAATGAGG 5477
      :  |||  |||||:::
448 5478
      :  |||  |||||:::
5478 TACTCTCTGCTACTGTGGCTGCTTCCATGTGGTGCCTCATCTGTGAC 5527
      :  |||  |||||:::
455 GlnAsnAsnAntYrIglYLeuAsnMetSerSerProProIsgIYSerPr 471
      :  |||  |||||:::
5528 AAGGTCCCTCTACATCTCTACATACCTGCAGCCCTTCTGGTCTCT 5577
      :  |||  |||||:::
471 ogIYLeuAlaProAsnIglYAsnIleMetIleSerProIArgAsn 486
      :  |||  |||||:::
5578 AATGTAGGT 5618
      :  |||  |||||:::
487 5578
      :  |||  |||||:::
5619 TCATTTTCAAAGCTCTCCTAAT 5665
      :  |||  |||||:::
501 AlagIYValHisSerProMetalIaSerSerGlyAsnThrGlyAsnHis 517
      :  |||  |||||:::
5666 GCC 5709
      :  |||  |||||:::
517 rPheSerSerSerLeuSerAlaLeuGlnAlaIleSerGlyIValG 534
      :  |||  |||||:::
5710 GACTAACCCCTGACAAATGCTGCACACCTACTACTACTTGGAGGTAG 5759
      :  |||  |||||:::
534 IYThrSerLeuLeuSerThrLeuSerSerProGlyProIYLeuAspAsn 550
      :  |||  |||||:::
5760 CTACTGTGTCTCTCTCAATGTCA 5806
      :  |||  |||||:::
551 SerProAsn 581
      :  |||  |||||:::
5807 GAACCAACTTCCCTGCTGCTGTGTTATGCACTGTGCTCCCAAGA 5856
      :  |||  |||||:::
561 sValSerAsnGlnAspSerIYSer 571
      :  |||  |||||:::
5857 GCCTTCTACTCAAGTAGCAACCACTGTGAGATACCACTCTCTCTC 5906
      :  |||  |||||:::
571 eugIYheYrYcysAspGlnAsnProValGlnSerSerMetYcysGlnSer 587
      :  |||  |||||:::
5907 TGCACAGACCTGAAGACTCAAAATCTCCCACTTCAATGTTGTTAA 5956
      :  |||  |||||:::
588 AsnSerAraGAspHisLeuSerAraIYsGlySerGlySerValG 604
      :  |||  |||||:::
5957 TTTCCAAACAA 6000
      :  |||  |||||:::
604 uGIYAla 619
      :  |||  |||||:::
6001 AGGAGCCCTTCTCTCCAGCCCAACAGAGACTCACACCAAGAAAGACC 6050
      :  |||  |||||:::
619 yLysLeuGlnLeuGlnLeuThrYcysSerSerAraIYsGlyHisSer 635
      :  |||  |||||:::
6051 CTACTGTATTAACGTTAGTCCAGGACCCCTTAAATTCCTCTCTTTC 6100
      :  |||  |||||:::
636 SerLeuThrAsn 645
      :  |||  |||||:::
6101 CAAAGTACATCTCTTCTCCAGAGATACCTTCTTCTCTCAACCACT 6150
      :  |||  |||||:::
645 rCysIYsGlnSer 652
      :  |||  |||||:::
6151 AGCAAAAGAAAGCTTGGGAGGCTCTCCATTAGTAAACAGCAGCAGA 6200
      :  |||  |||||:::
652 erValThrSerProSerGlyValSerSerThrSer 664
      :  |||  |||||:::
6201 GTATGACCTCCCTCTGGGTAACTCTCGGCTCTGTAAATCAAGACA 6250
      :  |||  |||||:::
665 6251
      :  |||  |||||:::
6251 GATTCTTATGACAGCCACAGACTCTGCTGCTCCCTTCTCAAAAGTCTCT 6300
      :  |||  |||||:::
677 uLeuGlnIYsHisArgIleuHisIYsLeuLeuGlnAsnGlyAsnS 694
      :  |||  |||||:::
6301 CATTTACCCCA 643
      :  |||  |||||:::
694 erProAlaGlyValAlaIYsIleThrAlaGlnAlaThrGlyAspThr 710
      :  |||  |||||:::
6344 6391
      :  |||  |||||:::
711 SerSerIleThrSerCysGlyAspGlyAsnVal 721
      :  |||  |||||:::
6392 ACTACAGCCAGCCCTTTCTACAGAGACACTCTCTTACTCTCTAAAA 6441
      :  |||  |||||:::
722 6442
      :  |||  |||||:::
6442 CCACCCAGTTAAAGAAAGTACTCTTACTACTTACCCTTGTTCACAG 6491
      :  |||  |||||:::
731 yLysGlnAsn 736
      :  |||  |||||:::
6492 CTTCAGAAATTTGCCCTGTGGCTCCATCCCCCAGAAATACCTGTGCTCC 6541
      :  |||  |||||:::
737 LeuLeuArgTyILeuLeuAspAraGAspProSerSerAraIaLeuSer 753
      :  |||  |||||:::
6542 CTGCTACCTTACTAGTGTGCTGGCCCTGGAATCCCAAGTCTGTGCCCTCAC 6591
      :  |||  |||||:::
753 sGlnLeuGlnPro 762
      :  |||  |||||:::
6592 CTCTCTTCCCCAGCTGGAGACTCTCCAGGTACAAAGAGTTGATGGTA 6641
      :  |||  |||||:::
762 alaSpAsnIYsMetSerGlnCysThrSerSerThrIleProSerSer 778
      :  |||  |||||:::
6642 TTTCTCAT 6676
      :  |||  |||||:::
779 GlnGluYsAspProIYsIleYsThrGlnThrSerGluGluYsSerG 795
      :  |||  |||||:::
6677 TCTCCCAA 6708
      :  |||  |||||:::
795 yAspLeuAspAsnLeuAspAlaIleLeuGlyAspLeuThrSerSerA 812
      :  |||  |||||:::
6709 T 6709
      :  |||  |||||:::
812 heTyIAsnAsnSerIleSerSerAsnGlySerHisLeuGlyThrIYsGln 828
      :  |||  |||||:::
6710 6745
      :  |||  |||||:::
829 GlnValPheGlnIYThrAsnSerLeuGlyLeuYsSerSerGlnSerA 845
      :  |||  |||||:::
6746 TACTAGCTGATTTCCCATCTCTTAAAGGTT 6789
      :  |||  |||||:::
845 IglYSerIleArgProProIYsAsnAraIaValSerLeuAspSerPro 862
      :  |||  |||||:::
6790 TCAGACTAAAGACCTCAACCAAGAAAGGT 6820
      :  |||  |||||:::
862 alSerValGlySerSerProProValYsAsnIleSerAlaPheProMet 878
      :  |||  |||||:::
6821 6865
      :  |||  |||||:::
879 LeuPro 894
      :  |||  |||||:::
6866 TCTCCAGTTGAAGCTTCAATTTCTTCCAGAGAAATAGTCTTTTCCAGG 6915

```

```

894 rGlnGlnAsnTyrGlySerSerMetGlyGlyProAsn..... 906
      :::::|||||
6916 CTCTAAAGACTCACCAGCCAGCAGCATCTCTCCATCTCCATCCCA 6965
907 .....ArgAsnVal 909
      :::::|||||
6966 AAGGGGCCCTACTCCCTCAGCTGTGACTCTCTCTCTCCCAAGAGATA 7015
910 ThrVal.....ThrGlnThrProSerSerGlyAspTyrGlyLeuProAs 924
      :::::|||||
7016 ACATCTACCCCAAGAGACCCCACTCTCTCAGTGTGAATCTGCCCTT 7065
924 nSerIleValArgMetGlyPro..... 932
      :::::|||||
7066 CCCCAGAGAGGCTCAGCTACTCTCAGCAGCCCAAGAGCTCTCTAT 7115
933 ..MetAsnSerAsnSer..MetGlyArgProGlyGlyAspTyr..... 945
      :::::|||||
7116 CCATGACTTCTCTCCCTCCCAAGAGGCGCGAGCACTCCAGCCCTATAA 7165
946 ..AsnThrSerIleuProArgProAlaLeuGlyGlySerIleProThrIleu. 961
      :::::|||||
7166 GGAATCCCAAGCTTCCCAATCCCAAGAGGCGCCCAACACCCCAAGCTGC 7215
962 .....ProLeuArgSerAsnSerIleProGlyAlaArg..... 972
      :::::|||||
7216 AACTCTCTCCCTCCCTTAAGAGGCGCCCACTTCCATCCCAAGATGG 7265
973 .....ProValLeu.....GlnGlnGlnGlnGln 981
      :::::|||||
7266 CCCCAGACCCCAAGCTGCACTCTCTCCCAAGAGAGGTCCACT 7315
981 t.....LeuGlnMetArgProG 987
      :::::|||||
7316 ACTCCATCCCAAGAGGCGCCCAACACCTCAGCTGCACTCTCTCTC 7365
987 LysIleIleProMetGlyMetGlyAlaAsnProTyrGlyGlnAlaAla 1003
      :::::|||||
7366 CCCCAGAGAGGCTCAGCTACTCTCATCCCAAGAGGCGCCCAACACCC 7415
1004 SerAsnGlnLeuGlySerTyrProAspGlyMetLeuSerMetGlnGln 1020
      :::::|||||
7416 CAGCTGCACTCTCTCTCCCAAG.....GAACTCAGCAGCT 7456
1020 LSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerIleuAsp 1037
      :::::|||||
7457 ACCCATCTCCCAAGAGGCGCATCA.....CACCCAGCTGCAC 7497
1037 euValGlyProProSerAsnLeuGln.....GlyGlnSerAspGln 1051
      :::::|||||
7498 TCCTCCCTCCCAAGAGATCCAGCAGCTACCCCACTCCCAAGAGGG 7547
1052 AlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThr 1068
      :::::|||||
7548 CCCCAGCAACCCCAAGCTGCA.....CTCTCTCCCTCCCAAGAGAGTCA 7594
1068 YLeuGlnGlnIleAspArgAlaLeuGlyIleProGlnLeuValAsnGln 1085
      :::::|||||
7595 GCTACCCCACTCTCAAGAGGCGCCCACTCTCCCAAGCTG..... 7634
1085 LysIleAlaLeuGlnProLysGlnAspAlaPheGlnGlyGlnAlaAla 1101
      :::::|||||
7635 ..CGACTCTCTCTCCCAAGAGAGCC..... 7661
1102 ValMetMetAspGlnLysAlaGlyLeuTyrGlyGlnThrTyrProAla 1118
      :::::|||||
7662 .....CAGCTACCCCAATCCCAAA 7681
1118 nGlyProPro..Met..... 1122
7682 GGGGCCCAAGCCCAAGCTGCAACTCTCTCTCCCAAGAGAGGTCT 7731

```

```

1123 .....GlnGlyGlyPheHisLeuGlnGlnGlnSerProse 1134
      :::::|||||
7732 AGCTACCCCAAGAGGCGCACCCAC.....AACCCAG 7769
1134 rPheAsnSerMetLeuAsnGlnMetAsnGlnGlnGlyAsnPheProLeuG 1151
      :::::|||||
7770 CTGCAGACTCTCTCTCCCAAGAGAGGTCTAGTACCCCAAGAGAGGTCT 7819
1151 LnglyMetHis...ProArgAlaAsnIleMetArgProArgThrAsn... 1165
      :::::|||||
7820 GGGGCCCAAGACACCCAGCTGCAACTCTCTCTCCCAAGAGAGGTCT 7869
1166 .....ThrProLys..... 1168
      :::::|||||
7870 AGCTACCCCAAGAGGCGCCCAACCAACCCAGCTCAACTCTCTC 7919
1169 .....GlnLeuArgMetGlnLeuGlnGlnArgLeuGlnGly 1181
      :::::|||||
7920 CTTCACCAAGAGAGGTCTAGTACCTACCCCAAGAGGCGCCCAACA 7969
1181 LnglnPheLeuAsnGlnSer.....ArgGlnAlaLeuGlnLeu 1193
      :::::|||||
7970 ACCCAGCTGCACTCTCTCTCCCAAGAGAGGTCTGCTACCCCAATC 8019
1194 LysMet.....GluAsnProThrAlaGlyGlyAlaAlaValMetAr 1207
      :::::|||||
8020 CCCCAGAGGCGCCCAACACCCAGCTGCAACTCTCTCTCCCAAGAG 8069
1207 g.....PrometMetGlnProGlnGlnGlnGlyPheLeu 1218
      :::::|||||
8070 GAGGCCAGCTACCCCAAGAGGCGCCCAAGAGGCGCCCAAGCTCA 8119
1218 snAlaGlnMetValAlaGlnArgSerArgGlnLeuLeuSerHisHisPhe 1234
      :::::|||||
8120 ACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8169
1235 ArgGlnGlnArgValAlaMetMetMetGlnGlnGlnGlnGlnGln 1251
      :::::|||||
8170 CCCCAGTCCCAAGAGGTCTACTCTCTCTCTCTCTCTCTCTCTCTCT 8219
1251 nGlnGlnGlnGlnGln.....GlnGlnGlnGlnGlnGlnGln 1264
      :::::|||||
8220 ACTTCACCCCAAGAGGCGCCCACTCTCTCTCTCTCTCTCTCTCTCT 8269
1264 LnglnGlnGlnThrGlnAlaPheSerProProProAsnValThrAlaSer 1280
      :::::|||||
8270 CCCCAGAGAGGCGCCAGGTACCCCAAGAGGCGCCCA..... 8301
1281 ProSerMetAspGlyLeuLeuAlaGlyProThrMetProGln..... 1294
      :::::|||||
8302 .....AAGGGGCCCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8330
1295 ..AlaProProGln..GlnPheProTyrGlnProAsnTyrGlyMetGly 1310
      :::::|||||
8331 CCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8380
1311 GlnProAspProAlaPheGlyArgValSerSerProProAsnAlaMet 1327
      :::::|||||
8381 TAAAGAGAGGCGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8425
1327 tSerSerArgMetGlyProSerGlnAsnProMetMetGlnHisProGln 1344
      :::::|||||
8426 .....CCAGCTGCAACCCCTC.....CCACAA 8450
1344 LaAlaSerIleTyrGlnSer..... 1351
      :::::|||||
8451 GAGAGTCCCGCTATGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8500
1352 .....GluMe 1353
8501 ATCTCCCAAGAGGAGCCCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 8550
1353 tLysGlyTyrProSerGlyAsnLeuAlaArgAsnSerSerPheSerGln 1370

```



```

526 LeuGlnAlaIleSerGlnGlyValGlyThrSerLeuSerThrLeuSe 542
      |||:|||||:
1194124TTTGAAGACCGTA.....TTTTTATTA..... 1194104

542 rSerProGlyProLysLeuAspAsnSerProAsnMetAsnIleThrGlnP 559
      |||:|||||:
1194103.....ATGAACGATTATATAC 1194088

559 roSerLysValSerAsnGlnAspSerLysSerProLeuGlyPheTyrcys 575
      |||:|||||:
1194087CAATCGAAATATACCCAGCGCTTTCCTTGAAGATTCATGTCATCA 1194038

576 AspGlnAsnProValGlnSerSerMetCysGlnSerAsnSerArgAspH 592
      |||:|||||:
1194037CAAACTCTGCCT..... 1194026

592 sLeuSerAspLysGlnSerLysGlnSerSerValGlnGlyAlaGlnAsnG 609
1194026..... 1194026

609 LnaArgGlyProLeuGlnSerLysGlyHisLysLysLeuGlnLeuLeu 625
      |||:|||||:
1194025.....CTTATTATCTAATA 1194011

626 ThrCysSerSerAspAspArgGlyHisSerSerLeuThrAsnProLe 642
      |||:|||||:
1194010.....CCGAGAGCTCACAAAGGCTTTCAGCCACTAATATTAT 1193976

642 uAspSerSerCysLysGlnSerSerValSerValThrSerProSerGly 659
1193975AAACAAACGCGCTTAACAGAGCCAGAAATTACTGTGTTGGC...TCCGCGAA 1193929

659 aLSerSerSerThrSerGlyGlyValSerSer..... 669
1193928TGTGAAAGACTAGAAAGACTGTGATTGTGCTTAAAGACAACTAAAGA 1193879

670 .....ThrSerAsnMetHisGlySerLeuLeuGlnG 680
1193878TTCATGATCCGCGAAACAAACAAATACCGCATCAAAACATCTATTAGA 1193829

680 uLysHisArgGlyLeuLeuHisLysLeuLeuGlnAsnGlyAsnSerProAlaG 697
      |||:|||||:
1193828AGAAAGCTAAATGATGACAGACAGATTTTAAAGAGGCCACAAA...TATA 1193782

697 LuValAlaLysIleThrAlaGlnAlaThrGly...LysAspThrSerSer 712
      |||:|||||:
1193781AACTACTATTCTGCACAGCAAGTGCACAGGCAATCAAGGAC..... 1193741

713 IleThrSerCysGlyAspGlyAsnValValLysGlnGlnGlnLeuSerP 729
      |||:|||||:
1193740TACTGACCTATGCGGAAATATGCTGTGTAAACGTAACACCTTTTACC 1193691

729 oLysLysLysGlnAsnAlaLeuLeuArgTyrlLeuLeuAspArgAspA 746
      |||:|||||:
1193690TGGTAAAGAA.....AATAAATTA.....TCGACG 1193665

746 SP.....ProSerAspAlaLeuSerLys 753
      |||:|||||:
1193664ACGCGGCATATCTGAAGAAGTCCGTAAGACCATCGGATCTTCAGAAAT 1193615

754 GlnLeuGlnProGlnValGlnGlyValAspAsn.....Lys 765
      |||:|||||:
1193614GATTCGATATTAAGCATAGCCGTAAGCATGATTCGAAATTAATTCGAA 1193565

765 sMetSerGlnCysThrSerSerThrIleProSerSerSerGlnGlnLysA 782
      |||:|||||:
1193564CGTTGACGATATTACT.....GAAAGAG 1193542

782 sPProLysIleLysThrGlnThrSerGlnGlnGly..... 793
      |||:|||||:
1193541AAAGTCTGCTGCTAAGCATATATGATGAGAAAGTAAGTAAGCAGCA 1193492

```

```

794 ....SerGlyAspLeuAspAsnLeuAspAlaIleLeuGlyAspLeuThrSe 809
      |||:|||||:
1193491GCTTAAGAGCATTTTGTGATTTTGCAGATGCTTGTGTATCTCAAGAGGTGC 1193442

809 rSerAspPheTyAsnAsnSerIleSer..... 818
      |||:|||||:
1193441CCAGCATCAAAATTAATCTCTATGTGACACCAAGTTACTCTAATAAAG 1193392

819 ..SerAsnGlySerHisLeuGlyThrLysGlnGlnValPheGlnGlyThr 834
      |||:|||||:
1193391CATCTCCCTCCCAACGGAAGTGTTTTAATCAACATCAAGTGCATCAT 1193342

835 AsnSerLeuGlyLeuLysSerSerGlnSerValGlnSerIleArgProP 851
      |||:|||||:
1193341TCACGCTTG...ATCGAAACCGAAGATCAAGAAAGAGCTTACCCCCC 1193295

851 oTyAsnArgAlaValSerLeuAspSerProValSerValGlySerSerP 868
      |||:|||||:
1193294ATTAAACTTTCAATCTTTGTGACGAA..... 1193267

868 roProValLysAsnIleSerAlaPheProMetLeuProLysGlnProMet 884
      |||:|||||:
1193266.....CTCAATCTCTTTGAAAAAAACATTAATACAGGATTTTACCGCTG 1193225

885 LeuGlyGly.....AsnProArgMetMetAspSerGlnGlnAs 897
      |||:|||||:
1193224TATTAATGTGATTTAATGACAGCGCCAAAAAAGATGATTTCTTACCANT 1193175

897 nTyGlySerSerMetGlyGlyProAsnArgAsnValThrValThrGln 914
      |||:|||||:
1193174TATTCCAATTTCCAA.....TCGGTGTGTGAC 1193146

914 hrProSerSerGlyAspTrpGlyLeuProAsnSerLysAlaGlyArgMet 930
      |||:|||||:
1193145TGTGATATATGATTTTAAATGATTTTAACTTGTGACGCTCAATGATTGACGAA 1193096

931 GlnProMetAsnSerAsnSerMetGlyArgProGlyLysArgTyAsnThr 947
      |||:|||||:
1193095GAACCAATCAATATCAACAAATGAGCAAAAGCGGGATGTGTTAT..... 1193051

947 rSerLeuProArgProAlaLeuGlyLysSerIleProThrLeuProLeuA 964
      |||:|||||:
1193050.....GGTATAGAAAGAAATCATTAATTTA 1193026

964 rGSerAsnSerIleProGlyAlaArgProValLeuGlnGlnGlnGln 980
      |||:|||||:
1193025GACACACAGACGATACCTTCATTTACGC...TATCTCAAAATTAACAACCCG 1192979

981 MetLeuGlnMetArgProGlyGlnIleProMetGlyMetGlyAlaAsnPr 997
      |||:|||||:
1192978ACTATATGCGCTGTGTAGAGCAAGATCAAGAACTGTCAAAAATTTATACAC 1192929

997 oTyGlyGlnAlaAlaIleSerAsnGlnLeuGlySerTrpProAspGly 1014
      |||:|||||:
1192928ATATATGCTAC.....AATGCG..... 1192913

1014 etLeuSerMetGlnGlnValSerHisGlyThrGlnAsnArgProLeuLeu 1030
      |||:|||||:
1192912.....GAATGATTAAGTGCACATATGACATATGCC...CTGCACA 1192877

1031 ArgAsnSerLeuAspAspLeuValGlyProProSerSerSerLeuGly 1045
      |||:|||||:
1192876TATATCTTAATTA.....TCAAAATTAATCAAGGTAG 1192848

1046 .....GlyLysSerAspGlnArgAlaLeuLeuAspArgGlnLeuIstHr 1060
      |||:|||||:
1192847AAGCCATGCGCAATGTTTGAAGCGTTTGTGA.....CAATTAAGTAAGAC 1192804

1060 euLeuSerAsnThrAspAlaThrGlyLeuGlnGlnIleAspArgAlaLeu 1076
      |||:|||||:
1192803GATTCATATTACTGACSTCAAGGCT.....CCCAAGACCAAT 1192766

1077 GlyIleProGlnLeuValAsnGlnGlnAlaLeuGlnLysProLysGlnAs 1093

```

```

1192765AATGCTCAACAACTGGTTGATTGAAAGCCCATTAATTTTCAGCAGACAG.. 1192718
      :::: :::: ::::: ::::: ::::: :::::
1192765AATGCTCAACAACTGGTTGATTGAAAGCCCATTAATTTTCAGCAGACAG.. 1192718
1093 pAlaPheInglInglInglAlaAlaValaIleMetLeuSpGlnLysAlaGlyL 1110
      ::::: :::::
1192717.....AATAGAAAGAAATATTCGC 1192702
1110 euTyGlyGlnThrTyProAlaInglTyProPromeGlnGlyGlyPhe 1126
      |||
1192701CTTTAGT.....GTAAATACAGAAATCCATCAAAAGGCGCAT 1192664
1127 HisLeuGlnGlyGlnSerProSerPheasnSerMetLeuasnMetas 1143
      :::: :::: |||::: ::::: :::::
1192663AGGAGACATCCGTTGGGCGCAGCATGTTTGAGCGCATTTAGAAAATGTAAGA 1192614
1143 nGlnGlnGlyasnPheProLeuGlnGlyMetHisProAlaAlaasnIleu 1160
      ::::: |||
1192613AAAAAAGAGAAAT.....ACGCCCGTCCCAATCCAA 1192582
1160 eArgProArg.....ThraasnThrPro 1167
1192581CCAGCGCAAGGAACCGTTAGATTGTAAAAATGAAAGTCCCTACCCCA 1192532
1168 Lys..... 1168
1192531GCAGAAATGCTCGTTTGAAAGCCCAAGACATGAGGCATTGAGAACGGA 1192482
1169 ...GlnLeuArgMetGlnLeuGlnGlnAlaArgLeuGlnGlyGlnPheL 1184
      ||||| |||::: |||::: |||:::
1192481TATACAAATTTGAGAAACAGACGTGAAAAAATGATTTCAACACAGCGCA... 1192436
1184 euAsnGlnSerArgGlnAlaLeuGlnLeuLysMetGlnLysasnProThrAla 1200
      ::::: |||::: |||:::
1192435..CAGCAAAATGACGACGACATTCATCCGAGTCAAGATCCAAATTCCT 1192388
1201 .....GlyGlyAlaAlaVala 1205
      ||| |||:::
1192387TCAAATGGCAATCTTCTTCAAAATTTGGCTAGAAATGGCGCAAGCTGGCG 1192338
1205 IMetArgPromeMetGlnProGlnGlnGlyPheLeuasnAlaIleMetV 1222
      ::::: |||||
1192337TCCAAAGACA.....AACCGAAACACTATATACGACAGATATCA 1192297
1222 aAlaGlnArgSerArgGlnLeuLeuSerHisHis..... 1233
1192267TTGAAAGTACTCGAGAGATTGTTGGAGCAAAAACAGACATTCGTCCT 1192247
1234 .....PheArgGlnI 1237
      ::::: |||
1192246GAGATGGCCCTTAAGCGCGCAAAAAACATAATTATGAAAGCTTGAGAGACA 1192197
1237 nArgValaIleMetLeuMetGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1254
      ::::: ||||| ||||| ||||| |||||
1192196ACAGACACAGCTTAAAGCAGCACCGATTTCAACACAAAGCAGCAGTTAC 1192147
1254 InGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnAla 1270
      ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::
1192146AGAGAGATCCGTCATGCGACGACATCGACAGCAACATACAAACAGGGCTCT 1192097
1271 PheSer.ProProPheasnValThrAlaSerProSerMetasPlyLeu. 1286
      ||| ||||| ||||| ::::: ||||| ||||| ||||| ||||| ||||| |||||
1192096TCAAGCTCGCGCACCGAAAA.....TCATCTCCATCTCAATCTTCTCTT 1192053
1287 .....IleuAlaGlyProThrMet 1292
      ::::: ||||| |||
1192052CCAAATATTTTCAACTTAATTCGCTCCAGATTAAGTACCAACACT... 1192006
1293 ProGlnAlaProGlnGlnPheProTyGln...ProAsnTyR 1306
      ||||| |||::: ||| ||||| |||||
1192005CCACAGGAGATTTTCCAAAGGTTCCAAAACAAATACACTTAATCTT 1191961
seq_name: /cgn2_6/prodata4/2/pna/us09_NEW_COMB_seq:us-09-335-032-12216

```

```

seq_documentation_block:
; Sequence 12216, Application US/0935032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; TITLE OF INVENTION: Transcriptome
; FILE REFERENCE: 01107,78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 12216
; LENGTH: 924430
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-12216

```

US-09-335-032-12216

[illegible]

**alignment\_block:**

US-09-041-994-2 X US-09-335-032-12216/rev .

Align seg 1/1 to reverse of: US-09-335-032-12216 from: 1 to: 924430

24 ThrProGlyGlnGlyLeuThrcysSerglyGluLysArgArgArgGlu 40

592258 ACCGAGGGTGAAGGT...GAAGGTGAAATGAAAAATAGGAAGAA 592

40 ngluserlvstvrllgglugluLeuA]agluLeuIleSerA]AsnLeu 57

592217 GGAAGAAAGTCTAGAGGA...C 592195

57 **erAsnT**]eAsnAsn**pheAsnVa**]f.vsp**ProAsn**.vs**CvSA**]at]eIeul.vs 73

502104 CAGGTTTTTACGATTCAGGACGACGATTCACCGACGACGATTAAGTACGT 502145

74 Clutha-vai-Arao] at] abao] at] at] weo] no] no] vt] wehxt] oeo 80

[illegible]

CC BY-NC-ND 4.0 International license.

4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
8

**P**

**K**

**P**

**K**

**—**

**— —**

**— —**

**· ·**

**· ·**

**· ·**

**—**

**— —**

**— —**

**P**

**· ·**

**· ·**

**· ·**

**K**

[illegible][illegible]

140 nvaatinnut yilbequtinnut yilbequtinnut sputtevaatinnut selv 131

392010 1A1ACCCAG1AAAC11GAI1AACAA1GGCGAC11AC1AC1AGACAC1GAA6 391901

```
13/ atty.....Asn11eleuhtSGluasparglysasprne 16;
```

591960 GTTACAAGCCAAACAAGATATGATGCATGCACACAGACATACAA..... 591917

170 LeuLysAsnLeuProLysSerThrValAsnGlyValSerTrpThrAsnG1 180

591916 CTGAAGCAGTTGTTAAAGGACGAATTCGGTAATGATTCAGATATTCCAG 591867



```

186 uProGlnArgGlnIynSerHisThrPheAsnGlySerMetLeuMetLysT 203
      :::::::::::::::::::: ||| ::::::::::::::::::::
591866 TTTTATCAAGGAGGACAA.....AATGGCAACATCAAGATGTGA 591826
      :::::::::::::::::::: ||| ::::::::::::::::::::
203 hrrProHisAspIleuGlnAspIleAsn.....AlaSerProGlu 216
      :::::::::::::::::::: ||| ::::::::::::::::::::
591825 ATCTTAGACACACTTACTGTATTTATCTCCCGACAGTACGCCCTCAACA 591776
      :::::::::::::::::::: ||| ::::::::::::::::::::
217 MetArgGlnArgTyrGlnIuThrMetGlnCysPheAlaLeuSerGlnProAr 233
      :::::::::::::::::::: ||| ::::::::::::::::::::
591775 ATGGCCACACAGCGGATCAAGTTTAGCGTTAGATGCA..... 591740
      :::::::::::::::::::: ||| ::::::::::::::::::::
233 galMetLeuGlnIuGlnIuGlnAspLeuGlnSerCysMetIleCysValA 250
      :::::::::::::::::::: ||| ::::::::::::::::::::
591739 .....ATGGATCCAGAGAGTACTTTGAAATTCGTTATCTTTCACCTT 591697
      :::::::::::::::::::: ||| ::::::::::::::::::::
250 IatArgArgIeThrThrGlnIuArgTyrPheProSerAsnProGluSer 266
      :::::::::::::::::::: ||| ::::::::::::::::::::
591696 CTGGTCCGATACCTAGTAGAGACGAACATCATTTTAGCAAAACAAACACAGT 591647
      :::::::::::::::::::: ||| ::::::::::::::::::::
267 PheIleThrArgHisAspLeuSerGly..... 275
      :::::::::::::::::::: ||| ::::::::::::::::::::
591646 AAAAGTGA AAAAACCTACAGCAAGAAAGAACAGACGAGCGGTGACGCCA 591597
      :::::::::::::::::::: ||| ::::::::::::::::::::
276 .....LysValIValAsnIleAspThrAsnSerLeuArgSer 287
      :::::::::::::::::::: ||| ::::::::::::::::::::
591596 GAAAGATACATACAGCAAGATTAAGATTCAGTATGACACAGTTAAAGTGC 591547
      :::::::::::::::::::: ||| ::::::::::::::::::::
288 ..SerMetArgProGlnPheGlnAspIleIleArgArgCysIleGlnArg 303
      :::::::::::::::::::: ||| ::::::::::::::::::::
591546 TAAATTTTAACTTATGCTGAAATGATCTG..... 591518
      :::::::::::::::::::: ||| ::::::::::::::::::::
304 PhePheSerLeuAsnAspArgIleGlnSerThrPserGlnIysArgHisTyrG 320
      :::::::::::::::::::: ||| ::::::::::::::::::::
591517 AAATTTTCCGTAATGATCACCGAAACCTATGTTACGTCAAATTAATTCGT 591468
      :::::::::::::::::::: ||| ::::::::::::::::::::
320 nGluAlaTyrLeuAsnGlnIHisAlaGlnThrProValTyrArgPheSerL 337
      :::::::::::::::::::: ||| ::::::::::::::::::::
591467 TTGGAGGTG.....GATGTCCCAATTCGT..ATCAGTG 591436
      :::::::::::::::::::: ||| ::::::::::::::::::::
337 euAlaAspGlyThrIleValThrAlaGlnThrLysSerLysleuPheArg 353
      :::::::::::::::::::: ||| ::::::::::::::::::::
591435 ATATGTGAAGGACACCTACCAAAATGCGAGCGCTTAGCGCATGTTCGCA 591386
      :::::::::::::::::::: ||| ::::::::::::::::::::
354 AsnProValThrAsnAspArgHisGlyPheValIleSerThrHisPheLeuG 370
      :::::::::::::::::::: ||| ::::::::::::::::::::
591385 ATGTAGGAAAAAGACTGACGCCACTTGGGTGTAGGCCAACTTTATTAGCGA 591336
      :::::::::::::::::::: ||| ::::::::::::::::::::
370 nArgGlnGlnAsnGlyTyrArgProAsnProAsnProValGlnGlnGlyI 387
      :::::::::::::::::::: ||| ::::::::::::::::::::
591335 GATCTCCAGGAATGCTAT..... 591317
      :::::::::::::::::::: ||| ::::::::::::::::::::
387 IeArgProProMetAlaClyCysAsnSerSerValGlyGlyMetSerMet 403
      :::::::::::::::::::: ||| ::::::::::::::::::::
591317 ..... 591317
      :::::::::::::::::::: ||| ::::::::::::::::::::
404 SerProAsnGlnGlyLeuGlnMetProSerSerArgAlaTyrGlyLeuAl 420
      :::::::::::::::::::: ||| ::::::::::::::::::::
591316 .....AATATCTCTATCTACTATCGAAGAAAGCTGGCAACAC 591279
      :::::::::::::::::::: ||| ::::::::::::::::::::
420 aAspProSerThrThrArgGlnMetSerSerGlyAlaArgTyrGlyGlySerS 437
      :::::::::::::::::::: ||| ::::::::::::::::::::
591278 TGTAT...TCCACGAGGAGATTATTTCGCATCANTGAACAGAAATGGCAGC. 591233
      :::::::::::::::::::: ||| ::::::::::::::::::::
437 eArgAsnIleAlaSerLeuThrProGlyProGlyMetClnSerPro..... 451
      :::::::::::::::::::: ||| ::::::::::::::::::::
591232 .....AACTACCAATGGGCGCTGTATTTATTAACACCCCGATAGA 591194
      :::::::::::::::::::: ||| ::::::::::::::::::::
451 ..... 451
      :::::::::::::::::::: ||| ::::::::::::::::::::
591193 ACGATGCTCGCTTAGCGGAGACTAATACTAAAAAAACCTGAACTCTT 591144

```

452 ..... SerSerGyr..... Glua 456

59113 TAAATCCGGTCTTAACAGCACAATAGACCGTTGTATTGTAAGACAGTG 591094

456 snAsnAsnTygIglyLeuasnMetSerSerProProlHisGlySerProGly 472  
||||| :||| :||| ||| :||| |||

591093 ATAACGAAGTCGATACAGAGAAAATAACACCAATTTTTTGC...GCC 591047

473 LeuAlaProAsnGlnGlnAsnIleMetIleSerProAlaGsnArgIySe 489  
::::: ||||| :||| :||| ||| :||| |||

591046 TTTCGT...AATAGCATTTACTGATCGTTATCTTACAGAAGCTGTGGGAT 5910000

489 rProlYstIleAlaSerHisGlnPheSer..... 498

590999 ACCT....ACCTCAAGAAATTTTCCAATAATACAGAGGGTAGGTTTC 590956

499 ..... ProValAlaGlyValHisSerProMetaIlaSerSer 510  
:::||||| ::|||

590955 ATATGCAATTATTGGACGTTAGCAGAGTTACAGAAAGCTCTATATTCTATAC 590906

511 GlysAsnThrGlyAsnHisSerPheSerSerSerSerLeuSerAlaLeuGI 527  
::||| :||| ||| ||| ||| :|||

590905 AATGAGCTGTGCGATCATTTCTTTCACCACAGTCAGC.....CTTA 590855

527 nAlaIleSerGlnGlyValGlyThrSerLeuSerThrIleSerSerP 544  
::::: :||| :||| ||||| :||| :||| |||

590864 TAGTCTCATCTTAAGAACATAACTAATCTTCATGGTCTCTGCC...TCCCCC 590818

544 roGlyProlYstLeuAsnSprasnsSerProtnMetAsnIleThrGlnProSer 560  
||| :||| ||||| ||| :||| ||||| :

590817 CTAAATAGAACGTTGGATAACTTT...GACTCAGAAATTTCTTCAGGTGCC 590771

561 LysVal.....SerAsnGlnAspSerLys..SerProLeuGlyPh 573  
::||| :||| ||||| :||| :||| :|||

590770 AAAACGCGATTATTAGAGCGCATCAGGAAGAAGAAATTCACAGACGTAATTT 590721

573 eTyfCyAsp.....GlnAsnProValGluSerSerMetC 585  
::||| :||| :||| :||| :|||

590720 TTGGAGAGACCCGTTAGTCGACATCGACAACTTATCGATATTATGACAATG 590671

585 ySgInSerAsnSerArgAspHisIleSerAsp..... 595  
::||| :||| :||| :||| :|||

590670 ATGATTTCTGATTAACATCGATGAAGATACGACGTATCACACAAAGCAAC 590621

596 ..LysGluSerLysGluSerSerValGlnGlyAlaGlnArgI 611  
::||| :||| :||| :||| :|||

590620 ATTATAGAAAAAGGCAATTCAGTCMAAACCCGCAAGCTACTATAAGC 590571

611 yProLeuGluSer...LysGlyHisLysLysLeuGlnLeuLeuThrc 627  
||| :||| :||| :||| :||| :||| :|||

590570 CCCCAGAAAGAAATGTGAGGGGAGCAGCAATAACAGCAAGTTTTCGCG 590521

627 ySSerSerAspArgGlnHisSerSerLeuThrAsnSerProLeuAsp 643  
::||| :||| :||| :||| :||| :|||

590520 CTTCCTCTGATGTAAAGAAATGCGTCGACCTG..... 590489

644 SerSerCysLysGluSerSerValSerValThrSerProSerGlyValSe 660  
::||| :||| :||| :||| :||| :|||

590488 .....GTGAG 590484

660 rSerSerThrSerGlyValSerSerThrSerAsnMetHisGlySerL 677  
||| :||| :||| :||| ||| |||

590483 TTCACATAGTACTCTCAGGA.....TCCACGGCCCAT 590453

677 euLeuGlnGlnLysHisArgIleuHisLysLeuLeuGlnAsnGlyAsn 693

590453 ..... 590453

694 SerProAlaGluValAlaLysIleThrAlaGlnAlaThrGlyLysSprH 710  
||| :||| :||| :||| :||| :|||

590452 .....AAATCTAC 590445

710 rSerSerIleThrSerCysGlyAspLysAsnValValLysGlnGlnGlnH 727





```

526728 AATTAACGGCACCCGACACTAGTAAATACATCGGCTCGAATTAACCTCC 526777
      ::::: :::: |||:::|||||:::
538 .....LeuSerThrLeuSerSerProGly...ProLysL 548
      |||:::|||||::: :::: |||:::
526778 ACCTCCTCTTTACCTTACCTTACCTTACCAACAAATAAGCCAAATA 526827
      |||:::|||||::: :::: |||:::
548 euAspAsnSerProAsnMetAsnIleThr..... 557
      :::: |||:::|||||:::
526828 TAATAGAGAGGCCCAATGACAGCTCAATTAATCAAGAAATACCTTTTA 526877
      :::: |||:::|||||:::
558 .....GlnProSerLysValSerAsnGlnAspSerLysSe 569
      :::: |||:::|||||:::
526878 GGTGAAACCTGTGTAGATGATACAAAGCGGAGAAATGCTCCGCGAATTC 526927
      :::: |||:::|||||:::
569 rProLeuGlyPheThrCysAspGlnAsnProValGluSerSerMetCysG 586
      |||:::|||||:::
526928 AACCAACACAC.....GATTAATGCTCCAGTAGCAAAATGATGGCTGC 526968
      |||:::|||||:::
586 InsertAsnSerArgAspHisLeuSerAspLysGluSer..... 598
      :::: |||:::|||||:::
526969 GTATACCGGATACCTCGAACGAGATGATTAATGAATAACAACAAATATG 527018
      :::: |||:::|||||:::
599 ...LysGluSerSerValGluGlyAlaGluAsnGlnArgGlyProLeuG 614
      |||:::|||||:::
527019 AAGAAGAAATAAACATTAATAGTGGGAAATGAACGTAATGATGACAC 527068
      |||:::|||||:::
614 uSerLys..... 616
      |||:::|||||:::
527069 CAGCAAAATATGACATACATCTACTAAACAGCGCTTCACGCGACCTT 527118
      |||:::|||||:::
617 ..GlyHisLysLysLeuGlnLeuLeuThrCysSer..... 628
      |||:::|||||:::
527119 TGGGCGTAGACAGACAACTACTAGCGCTCTCAGTGTAGTCTCCACAGCT 527168
      |||:::|||||:::
629 ...SerAspAspArgGlyHisSerSerLeuThrAsnSerProLeuAspSe 644
      |||:::|||||:::
527169 AATGCTGACAAATCACATTAATTAACAAGAAACACGACGACACACAAACA 527218
      |||:::|||||:::
644 rSerCysLysGluSerSerValSerValThrSerProSerGlyValSers 661
      :::: |||:::|||||:::
527219 CGGCAATTAACAGTAATATGTCATCCAAATAAACCAATGCCGATATCGACA 527268
      :::: |||:::|||||:::
661 eSerThrSerGlyValSerSerThrSerAsnMetHisGlySerLeu 677
      |||:::|||||:::
527269 ATTCTAAACCCCACTTGACGCTTCTTCACTCTTAAC...AATAATGCATAA 527315
      |||:::|||||:::
678 LeuGlnGluLysHisArgTyrIleLeuHisLysLeuGlnAsnGlyAsnSe 694
      :::: |||:::|||||:::
527316 AATGACGACTCTCAATGAG.....AGTAATTCAGAAATA 527347
      :::: |||:::|||||:::
694 rProAlaGluValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrS 711
      |||:::|||||:::
527348 ACCAACAAGGCGGATTTTTCGCTCGCAAGCGCTGCT..... 527384
      |||:::|||||:::
711 eSerIleLeuThrSerCysGlyAspGlyAsnValValLysGlnGlnLeu 727
      |||:::|||||:::
527385 .....ACACTGTAGGTGAATAAGTAATAGTCATCTCGAGGAACA 527426
      |||:::|||||:::
728 SerProLysLysLysGlnAsnAsnAlaLeuLeuArgTyrLeuLeuAspAr 744
      :::: |||:::|||||:::
527427 TTGTGTTATGAATCGGACAGTAATTCG...ACTAAACACCTAATATTTCC 527473
      :::: |||:::|||||:::
744 gAspAspProSerAspAlaLeuSerLysGlyLeuGlnPro..... 757
      |||:::|||||:::
527474 TGACGTCTCCAGCAGACAGACAGACAGACAACTCCAAACAAAC 527523
      |||:::|||||:::
758 .....GlnValGluGlyValAspAsnLysMetSerGlnCysThr 770
      |||:::|||||:::
527524 ACCAACAAGCAAAATAATGATGAATAACCTCAAAAGATA..... 527561
      |||:::|||||:::
771 SerSerThrIleProSerSerGlnGluLysAspProLysIleLysThr 787
      |||:::|||||:::

```

```

527562 AGCCCCCATTTGCTAAACAAATAACAAATAATTTTAAGCCGACTGAATA 527611
      :::: |||:::|||||:::
787 rGluThrSerGluGluGlySerGlyAspLeuAspAsnLeuAspAlaIleL 804
      :::: |||:::|||||:::
527612 TTCAAGACATATATAGCACGTGTGCCATATTTGAATAAC.....ACAA 527652
      :::: |||:::|||||:::
804 euGlyAspLeuThrSerSerAspPheThrAsnAsnSerIleSerSerAsn 820
      :::: |||:::|||||:::
527653 TCGCGACTATTAAGCACAAATCCGAACCTTGAAATTTAATGTGTGACGAC 527702
      :::: |||:::|||||:::
821 GlySerHisLeuGlyThrLysGlnGlnValIleGlnGlyThrAsnSerLe 837
      :::: |||:::|||||:::
527703 AATAACAATCTGATGTGCGGACACAAATCACCTGACGAGATTAAGCAGTAT 527752
      :::: |||:::|||||:::
837 uGlyLeuLysSerSerGlnSerValGlnSerIleArgProProTyrAsna 854
      :::: |||:::|||||:::
527753 AAACAGAGCCACCCGATCAATTCGACGACGACGACCA..... 527795
      :::: |||:::|||||:::
854 rGluValSerLeuAspSerProValSerValGlySerSerProProVal 870
      :::: |||:::|||||:::
527796 .....ATGGATGTACAAATCGTAGATTTGTAATACCTGTACAAACCGAC 527840
      :::: |||:::|||||:::
871 LysAsnIleSerAlaPheProMetLeuProLysGlnPro..... 883
      :::: |||:::|||||:::
527841 ACCAATGTTATGCC.....AGTCCGCTGATAAAGAGTC 527875
      :::: |||:::|||||:::
884 .....MetLeuGlyLysAsnProArgMetLeuAspSerGlnG 896
      :::: |||:::|||||:::
527876 AAGCTTAGATATCCCTATCTAAAGTTTCTCCACATTTACTTATCCACAT 527925
      :::: |||:::|||||:::
896 LuAsnTyrGlySerSerMetGlyGlyProAsnArgAsnValThrValThr 912
      :::: |||:::|||||:::
527926 CAAGCAACGGTAACAAATATCGTGTGCCAATGTGCCCAAAATTCGCGAG 527975
      :::: |||:::|||||:::
913 GlnThrProSerSerGlyAspTyrGlyLeuProAsnSerLysAlaGlyArg 929
      :::: |||:::|||||:::
527976 GAATTCGAA.....CCAAACAAATGATATTTTCAC 528004
      :::: |||:::|||||:::
929 gMetGluProMetAsnSerAsnSerMetGlyArgProGlyGlyAspTyrA 946
      :::: |||:::|||||:::
528005 GAAGAATATCTTTCCAAATTTACTTTGAGACATTCCTCTGCTAATATAGA 528054
      :::: |||:::|||||:::
946 snThrSerLeuProArgProAlaLeuGlyGlySerIleProThrLeuPro 962
      :::: |||:::|||||:::
528055 AT..... 528056
      :::: |||:::|||||:::
963 LeuArgSerAsnSerIleProGlyAlaArgProValLeuGlnGlnGln 979
      :::: |||:::|||||:::
528057 .....TCTAATTTATGCTGACAAACAAAGGCTCTTAGAACAAACAGTGT 528100
      :::: |||:::|||||:::
979 nGlnMetLeuGlnMetArgProGlyGluIleProMetGlyMetGlyAlaA 996
      :::: |||:::|||||:::
528101 AAAGATATTTGATTCAAACCTTAATGACGCTCTTTA..... 528137
      :::: |||:::|||||:::
996 snProTyrGlyGlnAlaAlaIleAsnAsnGlnLeuGlySerTyrProAsp 1012
      :::: |||:::|||||:::
528138 .....CGGAGATACCTGCGGTACCGGAT 528161
      :::: |||:::|||||:::
1013 GlnMetLeuSerMetGluGlnValSerHisGlyThrGlnAsnArgProLe 1029
      :::: |||:::|||||:::
528162 ...SACGTTAATCTAGAAAGATTAATCAGAACGCGCGAATTAATTCACAC 528208
      :::: |||:::|||||:::
1029 uLeuArgAsnSerLeu.....AspAspLeuValGlyPro..... 1040
      :::: |||:::|||||:::
528209 AATGCAAAATATAGTGTGAAGAGATGAATTTTAACACAGCAACAAATA 528258
      :::: |||:::|||||:::
1041 .....ProSerAsnLeuGluGlyGlnSerAspGluArgAlaLeuLeuAsp 1055
      :::: |||:::|||||:::
528259 ACTTCCCGCATGCTTTAAACCTTTATGCGATTAACAAATGTTATGTAAGAG 528308
      :::: |||:::|||||:::
1056 GlnLeuHisThrLeuLeuSerAsnThrAspAlaThrGlnGlyGlnGlu 1072
      :::: |||:::|||||:::
528309 GAA.....AATAATGCTGACTGCTAATGTAAATCGA... 528341
      :::: |||:::|||||:::

```





22109 .....GCATTAAACACAGTAAACAGTGCAGACGTGT 22143  
 813 Asn.....AsnSerIleSerSerAsnGlySerHisLeuGlyThrLysG 828  
 22144 ATCTGCACCAAAATGGTGTGAACATACGTGCAGCATGAATTAATTAATCTGCCG 22193  
 828 IngInValPheGlnGlyThrAsnSerLeuGlyLeuLysSerSerGlnSer 844  
 22194 TGACAGCTTTAAAGCGTGCATTCGTGATTAAGTCAGACACAAACCTAGT 22243  
 845 ValGlnSerIleArgProProTyrAsnArgAlaValSerLeuAspSerPr 861  
 22244 GGTACTATGTCAATGCTGATCGCAATTAAGCTCAACCATATGATGAATA 22293  
 861 oVal.....SerValGlySerSerProProValLysAsnIle. 873  
 22294 AGTTACAGCTCCGCAAAATATGCTTATAGTGTACACCAACACCAAGCTTAA 22343  
 874 .....SerAlaPheProMetLeuProLysGlnPro 883  
 22344 CACCAGCAGATGTTACAAATGACGACGACGACGATGCTTAAGACG 22393  
 884 MetLeuGlyGlyAsnProArgMetLeuSerGlnGluAsnTyrGly.. 899  
 22394 CAGTTAAACGGTAAATCATTAATTTAGACTAGCAGAAACAAATGCTAACAC 22443  
 900 .....SerSerMetGlyGlyProAsnArg..... 907  
 22444 TGCATTATGATGTTAACTCTCTTAATATGTCGCCAAACCAAAACCTTA 22493  
 908 .....AsnVal...ThrVal... 911  
 22494 AAGAACAAGTGGTCAACGACGACGACGCTGCCAAATGTTCAAACTTGTCT 22543  
 912 .....ThrGlnThrProSerSerGlyAspTyrGlyLeuProAsnSer. 925  
 22544 GATATGACCAAAACATTAACACCTGCATGAAGGCTTACGACGATAGCAT 22593  
 926 .....LysAlaGlyArgMetGluProMetAsnSerA 936  
 22594 TGCAGATACACGACGATTAAGCAGCTCA.....AACTACACAG 22644  
 936 snSerMetGlyArgProGlyGlyAspTyrAsnThrSerLeuProArgPro 952  
 22645 ATGCAATGCAAAACAAACAACTGACTACAC..... 22666  
 953 AlaLeuGlyGlySerIleProThrLeuProLeuArgSerAsnSerIlePr 969  
 22667 .....AGTGCAGTCAC 22677  
 969 oGlyAlaArgProValLeuGlnGlnGlnGlnMetLeuGlnMetArgP 986  
 22678 TGCAGCAAAACCAATCAT..... 22696  
 986 roGlyGluIleProMetGlyMetGlyAlaAsnProTyrGlyGlnAlaIa 1002  
 22697 .....GCTCAACACACTAGTCCATCAATGAATGCGCAA 22729  
 1003 AlaSerAsnGlnLeuGlySerTyrProAspGlyMetLeuSerMetGluG 1019  
 22730 GAAATTAATCAAGC.....AAAGCAACCAAGTACAGCTAAACACAA 22770  
 1019 nValSerHisGlyThrGlnAsnArgProLeuLeuArg..... 1031  
 22771 AGCGTTAAACGCTCAAGAAAC.....TTAAGAACTGCGCAAAACAA 22811  
 1032 .....AsnSerLeuAspAspLeuValGlyProProSer 1042  
 22812 ATGCGAAGCAACATTGTAACGCTTAATGACTTAAGTACGCTCAAAA 22861  
 1043 Asn.....LeuGlyGly..... 1046  
 22862 GATGACGTGAACGTCAAATGCAAGGTGCAACGATGTTAATGAAGTAAC 22911

1047 GlnSerAspGluArgAlaLeuLeuAspGlnLeuHisThrLeuSerA 1063  
 22912 ACAAGCACCAAAATTAATGCG.....GATGACTTAATATACCTATGACGA 22955  
 1063 snThrAspAlaThrGlyLeuGlnGluIleAspArgAlaLeuGlyIlePro 1079  
 22956 ACTTG.....AAAATGCTATTCAGATCAGAAATACG.....ATTAG 22993  
 1080 GluLeuValAsnGlnGlyGlnAlaLeuGluProLysGlnAspAlaPheG 1096  
 22994 CAAGGTGTTAACTTCATCATGATGCCGACGACGACGAAACGTAATGATAT.. 23041  
 1096 nGlyGlnGluAlaValMetMetLeuAspGlnLysAlaGlyLeuTyrGly 1113  
 23042 .....ACAATGCAAGTGCAGCAAGCTGAACAAATTTTAAT..... 23077  
 1113 InThrTyrProAlaGlnGlyProProMetGlnGlyGlyPheHisLeuGln 1129  
 23078 .....AAAGCACAGTCCA.....AATCTTCAAAAGAC 23107  
 1130 GlyGlnSerProSerPheAsnSerMetLeuAsnGlnGln..G 1145  
 23108 GGTGTGCAAACTGCGCTTGAATAATGTACACGCTGCTAAATAAGCAATTGA 23157  
 1145 nGlyAsnPheProLeuGlnGlyMetHisProArgAla...AsnIleMetA 1161  
 23158 CGGTATATCAAAATGTTGCGAACGCTTAAGACAACTGCGAAATGATGATTGA 23207  
 1161 rGProArgThrAsnThrProLysGlnLeuArgMetGlnLeuGlnGlnArg 1177  
 23208 ATTAACCTAACATCAATTATTAATATGCAAAAAGAAAGCAATTGAAATCACA 23257  
 1178 LeuGlnGlyGlnGlnPhe.....LeuAsnGln...SerArgGlnAl 1190  
 23258 ATTGAAGGTGCGACAAACGTTGACAGTGTAAATCAAGGCTCTACAAACGCG 23307  
 1190 AlaGluLeuLys.....MetGluAsnProThrAlaGly.....GlyA 1203  
 23308 ATCTGAATTAATTAATACGCAATGACGCAATGCAAAATGCTATTAATGATG 23357  
 1203 AlaAlaValMetArgProMetMetGlnProGlnGlnGlyPheLeuAsnAla 1219  
 23358 AAGCGCTACAAACAAACGCTTATGTTACTCAAAACCTTGAAACAACT 23407  
 1220 GlnMetValAlaGlnArgSerArgGluLeuLeuSerHisHisPheArgG 1236  
 23408 AAACAAACACGCAAAATACAGCAATGACGCTTAAAGCCATTTAAACAAATGC 23457  
 1236 nGlnArgValAlaMetMetMetGlnGlnGlnGlnGlnGlnGlnGlnGln 1253  
 23458 ACAAAAGAGGCAATTAAACAAATGCTGTACAAACATCGACTGCTGTCAG 23507  
 1253 IngInGlnGlnGlnGlnGln..... 1260  
 23508 AAGCAACAGTAATGACGAAAGCAACAAATGTTGATGACGCAATGAC 23557  
 1261 .....GlnGlnGlnGlnGln 1266  
 23558 AAATTACTCAAAAGTATTCAGATTAATGCGACACAAACAAACAAACAA 23607  
 1266 nGlnThrGlnAlaPheSerProProProAsnVal.....ThrAlaSerP 1281  
 23608 TTATACGTGATGATCAAGTCAAGTAATAAAGAGCGGACGACAAATATGCTGCA 23657  
 1281 roSerMetAspGlyLeuLeuAlaGlyProThrMetProGlnAlaProPro 1297  
 23658 CACTGCGCAAGGTAATTTATGATCAAACTACAAAGTCCAACTTTGATGATCG 23707  
 1298 GlnGlnPheProTyrGlnProAsnTyrGlyMetGlyGlnGlnProAspPr 1314  
 23708 ACTGTTATC.....AATCA 23721

[illegible]







```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (35)..(3690)
; US-09-654-935A-142
```

## alignment\_scores:

```
Quality: 215.50      Length: 1247
Ratio: 0.396          Gaps: 61
Percent Similarity: 43.625      Percent Identity: 20.529
```

## alignment\_block:

```
US-09-041-994-2 x US-09-654-935A-142 ..
```

```
Align seg 1/1 to: US-09-654-935A-142 from: 1 to: 3690
```

```
404 SerProAsnGlnGlyLeuGlnMetProSerSerArgAlaTyrGlyLeuAl 420
||| : : : : : : : : : : : : : : : : : : : : : : : :
127 TCCTCTCCGAGGTCATCGCTATTCCACCTGCTCAGAGTGGGGACAGCG 176
420 aasProSerThrThrGlyGlnMetSerGlyAlaTyrGlyGlySerS 437
: : : : : : : : : : : : : : : : : : : : : : : :
177 GGAAGTAAAGTCTGCGCTCA...TCACATGCAAGATAAATCCACAG 223
437 eRasnIleAlaSerLeuThrProGlyProGly...MetGlnSerProser 452
: : : : : : : : : : : : : : : : : : : : : : : :
224 ACCACCTCATCTCCAGCGCATGCTGAAAAGGTGATCTCATGCTCCCGT 273
453 SerTyrGlnAsnAsnAsnTyrGly...LeuAsnMetSerSerProPr 467
: : : : : : : : : : : : : : : : : : : : : : : :
274 CGCACATGAGATGAGGGCATGCGCACCATCTTTCACAGGACTTAGACCC 323
467 ohISglySerProGly.....LeuAlaProAsnGlnGlnAsnI 480
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
324 TCAT...TCCCTGTGTAACCCGACATGTTGGCCCTGTGATCATCAGCG 370
480 leMetIle...SerProArgAsnArg.....GlySerProLysIleAla 493
: : : : : : : : : : : : : : : : : : : : : : : :
371 TCCCTTAATGCTCCAGCAATAGGAGAAATGGTCCCAAGCTATGCGG 420
494 SerHisGlnPheSerProValAlaGlyValHisSerPromet..... 507
: : : : : : : : : : : : : : : : : : : : : : : :
421 TATATTGTAGAAACAACCTGCGCTCCACGCCACCCCAAGATGTGGA 470
508 .....AlaSerSerGlyAsnThrGlyAsnHisSerPheSerSerS 522
: : : : : : : : : : : : : : : : : : : : : : : :
471 TGGCAAGAGTGAATCTAGTATATTCAGGGGGGAGGACACAGACTCGG 520
522 eRLeuSerAlaLeuGlnAlaIleSerGlnGlyValGlyThrSerLeuLeu 538
: : : : : : : : : : : : : : : : : : : : : : : :
521 AGCCCTGGGAATACAAATCTCAGAGTAATGGAAGGGCATCCCTCTGAG 570
539 SerThrLeuSerSerProGlyProLysLeuAspAsnSerProAsnMetAs 555
||| : : : : : : : : : : : : : : : : : : : : : : : :
571 CCGGATTTAGCACTCTGCTATTCCATCCCAAACT...AACTAGAG 617
555 nIleThrGlnProSerLysValSerAsnGlnAspSerLysSerProLeu. 571
: : : : : : : : : : : : : : : : : : : : : : : :
618 CAGCTGAGTTTGGACCAACGCTCCAAACAAAGAGATGCTGGTCGCTGT 667
572 .....GlyPheTyrCys.....AspGlnAsn 578
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
668 ATTCTGAGGACCAAGATGCTACTGTGATCTGTGCACATGACTCTGGA 717
579 ProValGlnUserSerMetCysGlnSerAsn..... 588
: : : : : : : : : : : : : : : : : : : : : : : :
718 CATGATCTGGGAATCTGTGCATATAGCAGTATGGCTTTGGGAACCCAG 767
589 .....S 589
768 GCACAGCGTGATCATATGTTTTTTGTGGAAGAGCTCAGAAAAACCAAGGG 817
```

```
589 eRArgAspHisLeuSerAspLysGln...SerLysGlnUserSerValGlu 604
: : : : : : : : : : : : : : : : : : : : : : : :
818 ACCGCTCCAAATACCAGATAAATCCATATCAAGAAACATCTCTTTGAG 867
605 GlyAlaGlnAsnGlnArgGlyProLeuGlnUserLysGlyHisLysLe 621
||| : : : : : : : : : : : : : : : : : : : : : : : :
868 AAAGCA.....AGAAGCTCCCTGCGACCTCCCGGACAGACTCCCT 911
621 uLeuGlnLeuLeuThrCysSerSerAspArgGly..... 633
: : : : : : : : : : : : : : : : : : : : : : : :
912 CCGCAGATCTCCCAAGAGACAGCAGCTGCAAGCGGACGAGTGTCAACG 961
634 .....HisSer.....SerLeuThr... 638
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
962 AGACCTGATGCGCACACTCCAGCACTGCTGACAGCTGAGCTCCAGCG 1011
639 .....AsnSerProLeuAspSerSerCysLysGln..... 648
1012 AAAGTGGCAGCTGCGCTCCACAGAGCCCTGCACTGATGATGGAAGACC 1061
649 .....SerSerValSerValThrSerProSerGly 659
1062 CTGGCTGCCCGCTCCCGAGCCAGAGCAGATTAGTGTGCGACAGCA 1111
659 aLserSerSerThrSerGlyGlyAla..... 667
: : : : : : : : : : : : : : : : : : : : : : : :
1112 TGACTTCGCGCACCAACCCCAATGCTATCTCTGTGCGGGGCCAGCCA 1161
668 .....SerSerThrSerAsnMetHisGlySerLeu..... 677
1162 TCGCAGAGTGCACACAGAGCGCTCAAGTCACAGTACAGGACCCCTGGGG 1211
678 .....LeuGlnGlnLysHisArgIleLeuHisL 687
1212 TTATTACATTGACTACACAGCGCATGACAGAA..... 1242
687 ySLeuleuGlnAsnGlyAsnSerProAlaGluValAlaLysIleThrAla 703
: : : : : : : : : : : : : : : : : : : : : : : :
1243 .....GATCCGGGGGAC..... 1254
704 GlnAlaThrGlyLysAspThrSerSerIleThrSerCysGlyAspGlyAs 720
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1255 CCGGAGGGGGCTGTTCACACAGAGTGGGTGCCCACTGGGAACGGGCC 1304
720 nVal..... 721
1305 AGTCCGCCATGTCCAAGAGGCTCCAGACCAATGCCCAAGTGCCTCG 1354
722 .....ValLysGlnGlnLeuLeuSerProLysLysGlnAsn... 734
1355 GTGGTTCAGTCAACCAAAAGATCATGTCCACAGAAAGTACACAGACTC 1404
735 .....AsnAlaLe 737
1405 ATTCTCCATCCAGTGGGTATTCACGACGATGATACACACAGACACT 1454
737 u.....LeuArgTyrLeuLeuAspArgAspAspProSerAspAlaLeuS 752
: : : : : : : : : : : : : : : : : : : : : : : :
1455 CACCCCTGCTGCTGTGTTTAAATCATGTCTCACACCAAAACGGGAAG 1504
752 eRlySGluLeuGlnProGlnValGlnGlyValAspAsnLysMetSerGln 768
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1505 GGAAG.....CCCAAGCCCAAGGTACCAAGAAAGAAAGTCTCTGTG 1545
769 CysThrSerSerThrIleProSerSerSerGlnLysAspProLysI 785
: : : : : : : : : : : : : : : : : : : : : : : :
1546 ATACTTTCAGTATCATCTTCTCATGCTCC.....ACTTCTCT 1583
785 eLysThrGlnThrSerGlnGlnGlySerGlyAspLeuAspAsnLeuAspA 802
: : : : : : : : : : : : : : : : : : : : : : : :
1584 TTCTTCTAGTACTTCTACTGAAGAGAGTGCACATATGAAAGCTGATC 1633
```





```

694 SerProAlaGluValAlaLysIleThrAlaGlnAlaThrGlyLysAspThr 710
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
235276 AGGCCAATAATATATGAA.....TGCTAATATGAATCTTAACCCAC 235315
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
710 rSerSerIleThrSerCysGlyAspGlyAsnValVal.LysGlnGln 726
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
235316 AGATGTCCTTGATCATCAGCAGCTCAGCAAGGCAACAGGTGGCCACAA 235365
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
727 LeuSerProLysLeuLysGlyAsnAsnAlaLeuLeuArgTyrLeuLeuAs 743
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
235366 TTAAAGAAATCAACACACAA..... 235386
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
743 PArgAspAspProSerAspAlaLeuSerLysGluLeuGlnProGlnValG 760
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
235387 .CAACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 235435
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
760 LngLysAlaAspAsnLysMetSerGlnCysThrSerSerThrIleProser 776
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
235436 AACAAATTAGTGAACGAGATGAAA.....GTGCACCTATTCCCAAA 235476
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
777 SerSerGlnGluLysAspPro..... 783
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
235477 CAATTACTGCAAGAAATTCCTTAACATTCCACCAATATCAACACCTGGCA 235526
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
784 .....LysIleLysThrGluThrSerGlnG 792
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
235527 GCAGGTCACTGCTTGTGCTCAACAAAGCTATTGACACCTCAAGAAATG 235576
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
792 LngLysSerGlyAspLeuAspAsnLeuAspAlaIleLeu..... 804
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
235577 AACCTCGAAGGAGTCTACAAAGATTCCACGACGATTCCTATTCAAGCA 235626
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
805 .....GlyAspLeuThrSerSerAspMetTyrAs 814
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
235627 AGGCTACAGCAACACACAGCAGCTCAAGCAGCAAGCTAATTAACAA 235676
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
814 nAsnSerIleSerSerAsnGlySerHisLeuGlyThrLysGlnGlnValP 831
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
235677 CAACGGGCTCCGCCAAATGTGTAATTT..... 235704
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
831 heGlnGlyThrAsnSerLeuGlyLeuLysSerSerGlnSerValGlnSer 847
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
235705 .....AACATTAACATTAATATTCCTCAACAGCAGCAAAATGCAA... 235743
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
848 IleArgProProTyrAsnArgAlaValSer.....LeuAspSerProVa 862
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
235744 .....CTCCCAATTCAAGTGGCAACACACCTTTGCAACAGCAATC 235787
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
862 LserValGlySerSerProProValLysAsn.....IleSerAlaPheP 877
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
235788 ATCACAAAATACCGTACCAACGCTCCTCAACCAAAATTAACCAAAATCTT 235837
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
877 roMetLeuProLysGlnProMetLeuGlyGlyAsnProArgMetMetAsp 893
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
235838 CTCACAGAGGACGACGACGCTTATTA.....CAAGACGATCGAA 235878
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
894 SerGlnGluAsnTyr.....GlySerSerMetGlyLysP 905
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
235879 ACCTGCACAAATTTGAAAAACACAAATGGGATGATGACGCGAACC 235928
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
905 oAsnArgAsnValThrValThrGlnThrProSerSerGlyAspTyrGly 922
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
235929 TGTCAAGCAAAAGTTTATTAGGAAA.....T 235954
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
922 euProAsnSerLysAlaGlyArg.....MetGluProMet 933
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
235955 ACATTAAACCAAAAGCCCTGAGAAATAATCCAACTTTGAGAGATGTTAAG 236004
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
934 AsnSerAsnSerMetGlyArgProGlyGlyAspTyrAsnThrSerLeuP 950
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
236005 AACACAAATTAACGCTTAACAAACAGGCTGCAACCTTCAG..... 236043
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
950 oArgProAlaLeuGlyGlySerIleProThrIleuProLeuArgSerAsnS 967
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

236044 ..... 236053
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
967 erIleProGlyAlaArgProValLeuGlnGlnGlnGlnMetLeuGln 983
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
236054 ATGTCCCT.....ATGAATATCATCCAAACACAAACAA.....CAG 236091
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
984 MetArgProGlyGluIleProMetGlyMetGlyAlaAsnProTyrGly 1000
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
236092 AACACGAAACAAATTAATGACACACATTCGCTGCTGACACCT.....AA 236135
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1000 nAlaAlaAlaSerAsnGlnLeuGlySerThrProAspGlyMetLeuSerM 1017
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
236136 TGCTGCCGCTTCTCTCCACCAACAGAACCAAGTTTAATTATATACGA 236185
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1017 erGlnGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSer 1033
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
236186 TGCACACACAGCAACAGCTCAAGCTCAA..... 236214
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1034 LeuAspAspLeuValGlyProProSerAsnLeuGlnGlyGlnSerAspG 1050
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
236215 .....GCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGC 236246
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1050 uArgAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaT 1067
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
236247 ACAAGCTCAAGCAGCAGCGGCAACAGCGGCGCAAGCGCAAGCAAGCA 236295
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1067 hrcGlyLeuGlnGluIleAspArgAlaLeuGlyTleProGluLeuValAsn 1083
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
236295 ..... 236295
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1084 GlnGlyGlnAlaLeuGlnProLysGlnAspAlaPheGlnGlyGln..G 1099
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
236296 CAAGCACAAGCACAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 236345
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1099 uAlaAlaValMetMetAspGlnLysAlaGlyLeuTyrGlyGlnThrTyrP 1116
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
236346 GGCCCAAGCCCAAGCCCAAGCACAACAGCAGCAGCAGCAGCAGCAGC 236395
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1116 roAlaGlnGlyProProMetGlnGlyGlyPheHisLeuGlnGlnSer 1132
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
236396 CCCCCAACACACA..... 236409
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1133 ProSerPheAsnSerMetMetAsnGlnMetAsnGlnGlnLysAsnPheP 1149
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
236410 .....CAACAAGCTCAGCAGCAACACCTAAC...CC 236435
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1149 oLeuGlnGlyMetHisProArgAla..... 1157
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
236436 ACTACATGGGTGTACACCTACTGCAAAAGATGCGAAGTAATTAAGCAAT 236485
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1158 .....AsnIleMetArg 1161
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
236486 TGCCTTGGATGCTTCTTAAGACCAACCTTAAGCTTACGAGCGTAACAAT 236535
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1162 ProArgThrAsnThrProLysGln..LeuArgMetGlnLeuGlnGlnArg 1177
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
236536 TCTTTATCCAAATGAAGAAAGAAATTAATAATGAAGTAAAG..... 236580
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1177 gLeuGlnGlyGlnGln..PheLeuAsnGlnSerArgGlnAlaLeuGlnL 1193
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
236581 ....CAAGTCAAAAAGCTTTTGTTCAGGTGAGTATTTTCCGCCCAAG 236626
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1193 eu.....LysMetGluAsn..... 1197
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
236627 TCTACATCATCACAAAGATGAATACTTCTTGAGAGAACTTTTTCGTTA 236676
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1198 .....ProThrAlaGlyAlaAlaVala 1205
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
236677 AGAATCTTTGTAAGAGATCTCAGAAATAATGCGCAGAGGTATATTTGT 236726
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1205 lMetArg..... 1207
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :

```

[illegible]

```

|||||
3441 TTCACGAAAGGGCCAGAAAGACCAAGGAAACCTCTCTACCTTCT 3392
507 MetalaserSerGlyAsnThrGlyAsnHisSerPheSerSerSer 523
      |||||
3391 CTGACCAAGGACGCGGACCAACCAACGCGCTGCTCAGCATACCT 3342
523 uSerAlaLeuGlnAlaIleSerGlyValGlyThrSerLeu...LeuS 539
      |||||
3341 TAAAGCCCGGAGTACCAAGCACCAAGCAAGCATCAGCACCCAGG 3292
539 eThrIleuSerSerPro.....GlyProIleu 548
      |||||
3291 GATCTTTTCTTCCAGCAAGGCGGAGGACCAAGGGGACCAAACTT 3242
549 AspAsnSerProAsnMetAsnIleThrGlnProSerIleValSerAsnG 565
      |||||
3241 .....TCACCA.....AGACGTCCAGCA 3222
565 nAspSerLysSerProLeuGlyPheTyrCysAspGlnAsnProValGluS 582
      |||||
3221 GGGCCAGTCTTCCACACGCGGACCTTTGC.....C 3190
582 eSerMetCysGlnSerAsnSerArgAspHisLeuSerAspLysGluSer 598
      |||||
3189 GCCTTCTTTCAGCCAGAACCCAGAGGCGCACGGGCTCCACATTTT 3140
599 LysGluSerSerValGly.....AlaGluAsnGlnArg 610
      |||||
3139 CCCAGAGGCGCCAGAAAGGACCGGANTTTCCGCGCAGCAGCACCAAGGA 3090
610 gGlyProLeuGluSerLysGlyHisLysLysLeu.....621
      |||||
3089 AACCCAGATGACATCCAGGGGACCAAGCGGTTCCGCGAAGACCTTTT 3040
622 .....LeuGlnLeuThrCysSer.....628
3039 GGCCTCCAGGACCAACCAATTACCAATGGGCGCAGGGGCTCATTC 2990
629 .....SerAspArgGly.....HisSerLeuThrAsnSe 640
      |||||
2989 GGGTCCCGGCAAGGGGCAAGGAGCAACAGCATCGCTTTAAGCACCAA 2940
640 rProLeuAspSerSerCysLysGluSerSerValSerValThrSerProS 657
      |||||
2939 GCATCACCAAGTGTCTCTTACCAACAGGTTGGCGCTCAGAACCAAG 2890
657 eArgValSerSerSerThrSerGlyGlyValSerSerThrSer...As 672
      |||||
2889 GGGGGGCGCAGAAAGCCAGAGGCGCGGAGTAAACCAAGCTTACCAAGC 2840
672 nMetHisGlySerLeuLeuGlnGlyLysHisArgIleLeuHisLysLeuL 689
      |||||
2839 GGTCTCCGGGGG.....CACCAACAA 2817
689 eugInAsnGlyAsnSerProAlaGluValAlaLysIleThrAlaGlnAla 705
      |||||
2816 GCTCCAGTGGGACCAAGGCGCGCTGG..... 2785
706 ThrGlyLysAspThrSerSerIleThrSerCysGlyAspGlyAsnValVa 722
      |||||
2784 .....ACCAACTTTTTCACCTTTGTGTACCAAGGCGCAC..... 2752
722 LysGlnGlnGlnLeuSerProLysLysLysGluAsnAsnAlaLeuLeuA 739
      |||||
2751 .....CAAGCAGGCGCAAGAGAGACCAATGGGCGGCTCCAGA 2712
739 rGlyr.....LeuLeuAspArgAspAspProSerAspAlaLeuSer 752
      |||||
2711 CCCAGGAGCCCAATCTTTTTCAGAGAGAGCATCAGCAGCCTTTGG 2662
753 LysGluLeuGlnProGlnValGlyValAlaAspAsnLysMetSerGlnCy 769
      |||||

```

```

2661 GACCACATACCTTTTGTACCTTTTAGCCCTTGAAGACCAAGCGTG 2612
769 s.....ThrSerSerThrIleProSerSerSerGlnLysAspProL 784
      |||||
2611 CCACACGAGTTTCACACAGCAATCCCTTGAAGGCGCCAGGGAG...CCCT 2565
784 ystIleLysThrGlnThrSerGluGlnGlySerGlyAspLeuAspAsnLeu 800
      |||||
2564 GCGTTACCGGAGCTCCAGGGGCGCACCAAGCATCCCTTAGCACCATT 2515
801 .....AspAlaIleLeuGlyAspLeuThrSerSerAspPheTyrAs 814
      |||||
2514 CGTTCGGGGAAGCAACCGTTGGCCCTCGGGGACCAAGACCAAGCAGG 2465
814 nAsnSerIleSerSerAsnGlySerHisLeuGlyThrLysGlnGlnValP 831
      |||||
2464 GGAACCTTCACACCAAGCGCTTG.....GCCAGGAAAACTCT 2424
831 heGlnGlyThrAsnSerLeuGlyLeuLysSerSerSerGlnSerValGlnSer 847
      |||||
2423 TCTGGCTCTTCTCCAGAGGGGCGACGAGGCAAGTCTTCCAAGG 2374
848 IleArgProProTyrAsnArg.....AlaValSerLeuAspSerPr 861
      |||||
2373 AACCAACCTTGTTCACCAAGTTTTGCCTGCTTCACCTGAGAGACCC 2324
861 oValSerValGlySerSerProValLysAsnIleSer.....874
      |||||
2323 AGCAAGACCAAGGGAAGACCTTGAATCCGGGGAAGCAGCAGCAAGGCGC 2274
875 ..AlaPheProMetLeuProLysGlnProMetLeuGlnGlyLysAsnProArg 890
      |||||
2273 CTGTTCACCTCTTCTCCAGAGGGAACGACAGCGGC..... 2236
891 MetMetAspSerGlnGluAsnTyrGlySerSerMetGlyLysProAsnArg 907
      |||||
2235 .....CAAGGGGTCC..... 2224
907 gAsnValThrValThrGlnThrProSerSerGlyAspTrpGlyLeuProA 924
      |||||
2223 .....TTAAGTCCAAAGCT.....CTCCAT 2202
924 snSerLysAlaGly.ArgMetGluProMetAsnSerAsnSerMetGlyArg 940
      |||||
2201 CTTTCCAGCAGGAGCAGCAGCGCCA.....GGGG 2170
940 gProGlyGlyAspTyrAsnThrSerLeuProArgProAlaLeuGlyLys 957
      |||||
2169 TCCGGGAAAAAACCTTGGGCTTCCAAAGCCTTGGCG.....GGGC 2129
957 eTlleProThrLeuProLeuArgSerAsnSerIleProGlyAlaArgPro 973
      |||||
2128 TTCTCCAGAGGCACTTTTAAGTCCAGGGAATCCCC..... 2091
974 ValLeuGlnGlnGlnGlnMetLeuGlnMetArgProGlyGluIlePr 990
      |||||
2090 .....ATCACCAACCAAGCT..... 2076
990 oMetGlyMetGlyAlaAsnProTyrGlyGlnAlaAlaIleAsnAsnGlnL 1007
      |||||
2075 .....GAACCAAGGCGCAC..... 2061
1007 eugLysSerTrpProAspGlyMetLeuSerMetGluGlnValSerHisGly 1023
      |||||
2060 .....AGTGCGCTGGGGGTTCGCGGCGGACCATTTTACCGCGCGG 2016
1024 ThrGlnAsnArgProLeuLeuArgAsnSerLeuAspAspLeuValGlyPr 1040
      |||||
2015 GACCAGGCGGCGCAGTTTTC..... 1995
1040 oProSerAsnLeuGlnGlnSerAspGluArgAlaLeuLeuAspLnt 1057
      |||||
1994 .CATCA.....GGACCAAGAAATGCCAGGCT..... 1968

```



```

1057 euH1sthrLeuSerAsnThrAspAlaThrGlyLeuGluGluLeasp 1073
1067 .....TCCATTTAAACCCCTTGGCCACCAG..... 1942
1074 ArgAlaLeuGlyLeProGluLeuValAsnGlnGlnAlaLeuGluPr 1090
1084 .....GGCAGACCAAGCTTCCACCGGAGGACCAAGCTTCCAC.. 1903
1090 GlysGlnAspAlaPheGlnGlnGlnGlnAlaValMetMetaspGlnL 1107
1092 ..CAAGAGATCCCTTTGGGGCCAGAGGGC..... 1873
1107 ysaIaGlyLeuTyGlyGlnThrTyProAlaGlnGlyProPrometGln 1123
1172 .....CAA 1870
1124 GlyGlyPheH1sLeuGlnGlnGlnSerProSerPheAsnSerMetMetas 1140
1169 GGAGAAACCCACGGTTCCACACCGCGGAGCCCTTTGGGACCAAGCAAC 1820
1140 nglMetAsnGlnGlnGlnAsnPheProLeuGlnGlyMetHisProArga 1157
1819 CACCATCTGCTGCCAGGGAAC.....CCACGG 1791
1157 laasnIleMetArgProArgThrAsnThrProLysGlnLeuArgMetGln 1173
1790 CT.....ACGAGTTCCACCAACGGCTCCGGCAGGG 1759
1174 LeuGlnGlnArgLeuGlnGlnGlnPheLeuAsnGlnSerArgGlnAl 1190
1758 GTCCGGGACAGCCCAAGTGGTCCGGGGTTCACTCGAGCTCTGGGCTTT 1709
1190 aLeuGlnLeuLysMetGlu.....AsnProThrAlaGlyG 1202
1708 CCTTCCTCTCCAGCAGGCGCCAGGGGGCTCTTGAACCAACAGAGGCC 1659
1202 LyAlaAlaValMetArgPrometMet.....GlnProGlnGln 1214
1658 AGGCTCTCCCTTTAGCACCAAGTGTCTCTTGCATGCCAGACACCA 1609
1215 GlyPhe..... 1216
1608 GGTCTCTCCGCTTGTACCTTGGAGACAGAGGCGCCCGGCGCTTG 1559
1217 .....LeuAsnAlaGlnMetValAlaGlnArgSerArgGluL 1229
1558 GGTCCAAAGGGGCGCTCGGCACCAAGCAAGCAAGCAAGCAAT 1509
1229 euleuSerH1sH1sPheArgGlnGlnArgValAlaMetMetMetGlnGln 1245
1508 ACCGAGGACCACTATTTGGCACCCTTACCAACCAAGCTGTCATCAGCAC 1459
1246 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1262
1458 CAGGGTTTCCAGCAGGCGCCAGCAGCAGCAGGCGGCGGCGCAGG 1409
1262 lngGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1274
1408 CTCACCCAGCGCACACCCCTGGGACCTTCAAGAGCTCGGGGCGCTTG 1359
1275 .....ProasnValThrAlaSerProSerMe 1283
1358 GGGAGCAGGCTTACCCCTTTAGCACCAACAGCAAGCAGGAGAGCCAGGA 1309
1283 taspGlyLeuLeuAlaGlyProThrMetProGlnAlaProGln... 1298
1308 GGACCA.....GGGGGGCGGTGGAGCAGAGGGGGCGGCGGACCA 1266
1299 .....GlnPheProTyGlnProAsnTyGlyMetGlyGlnGln 1311
1365 AGTAGCACTTATCATTTCCAGCAGACCAACAGGAGGGCGGAGGG..... 1222

```

```

1312 ProAspProAlaPheGlyArgValSerSerProProAsnAlaMetMetse 1328
1221 .....GCTCCAGGGGGCAGCCTCTCTCTCACCAGCAGGAGGCGCAGGG 1181
1328 rSerArgMetGlyProSer..... 1334
1180 G.....GGCCATCTTGACCAAGAGAGCTCCATTTTTCACCAGGAGG 1140
1335 .....GlnAsnProMetMetGlnHis..ProGlnAla 1344
1139 TGCCAGAGTTCACCCCTTAGCAGCAAGCAGCAGCAACATCTTCCCTTGGA 1090
1345 AlaSerIleTyGlnSerSerGlu.....MetLysG 1355
1089 CCATCC.....AAACCACTGAACACCTTGTGTCTTCCCTTATTCACAG 1046
1355 yTrpProSerGlyAsnLeuAlaArgAsnSerPhe...SerGlnGln 1371
1045 GAGGCGCAGCTGTTCGGGCAATCTCGAGCAGCAGCAGCAGGCGGAGGCG 996
1371 lnpheAlaHisGln.....GlyAsnProAlaValTySerMetVal 1384
995 CCACGGCTCACACGAGCAGCAGGAGTTTCCAGCTTCCCTCCATCATCTCTC 946
1385 HisMet.....AsnGlySerSerGlyHisMetGlyGlnMetMetas 1399
945 CATCTTTTCCAGGGGAGCCTGGGGGAGCTTGGGGA..... 910
1399 nPrometPrometSerGlyMetPrometGlyProasp 1411
909 .....CCCATGGAGCTCGAA 895

```

```
seq_name: /cgn2_6/prodata/2/pna/US09_NEW_COMB.seq:US-09-543-679A-3003
```

```

seq_documentation_block:
; Sequence 3003, Application US/09543679A
; GENERAL INFORMATION:
; APPLICANT: NYCE, Jonathan W.
; TITLE OF INVENTION: LOW ADENOSINE AMTI-SENSE OLIGONUCLEOTIDE,
; COMPOSITIONS, KIT & METHOD FOR TREATMENT
; OF AIRWAY DISORDERS ASSOCIATED WITH
; BRONCHOCONSTRICITION, LUNG INFLAMMATION,
; NUMBER OF SEQUENCES: 3111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: EPIGENESIS PHARMACEUTICALS, INC.
; STREET: 7 Clarke Drive
; CITY: Cranbury
; STATE: NJ
; COUNTRY: USA
; ZIP: 08512
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-R
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: N/A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/543,679A
; FILING DATE: 13-Apr-2000
; CLASSIFICATION: UNKNOWN
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/127,958
; FILING DATE: 1998-08-03
; ATTORNEY/AGENT INFORMATION:
; NAME: Amzel, Viviana
; REGISTRATION NUMBER: 30,930
; REFERENCE/DOCKET NUMBER: EPI-0067191b
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 609-409-3035
; TELEFAX: 413-254-9245
; TELEX: <unknown>
; INFORMATION FOR SEQ ID NO: 3003:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35459 base pairs

```





442	nmetseserprprohni	sglyserproglyleu	laiproasn	gln	la	479
338069	TTTTCTCTTACAAAGCAT.....					338041
479	snllemetleserpro	adasnarg	glyserprolys	llealaser	his	495
338040	ACAATTATACCTCAAGCCAGAAATTAACAAGAG.....					338009
486	glnpneserpro	vala	la	gly	his	512
338008	.....GTACAAACCCAGGCTCATGACA.....					337988
512	ntrhglyasn	his	serp	hes	ser	529
337984	.ACCGGGAACACT.....					337973
529	leser	g	l	g	l	545
337972	.....GTTATCTCGACAAAGGGC					337955
546	prolysleu	aspa	ns	ser	pro	562
337954	AAGCAGTACACATCCAAAGCCTTTTCAAACACTACATCGGGCCACAGC					337905
562	lserasn	g	l	as	ser	579
337904	TGTTTCTGAGACTACTAAACATGAGCCT.....					337876
579	roval	g	l	as	ser	595
337875	CTCAGAGATTCA.....AATGAT					337859
596	lys	g	l	as	ser	612
337858	AACATATATGAAGATTAAACACACAGGCGCAACAGCCGAGGAGAAA					337809
612	oleu	l	as	ser	l	629
337808	TCAAGAAAAAGACGAAGAGACCAACACAGCAG.....					337775
629	eraspa	arg	l	his	ser	645
337774	.....GAAGGCGATATATAC.....					337760
646	cys	l	as	ser	l	662
337759	..AAGAGAAACACAAACAAATAGAGCAACCTCTTATCTTCAAGAA					337713
662	trh	ser	g	l	as	676
337712	AACACCTCTAGACATCTGCTTCAACACCAAGAAATGCTGGGCTG					337663
676	erleu	leu	g	l	as	692
337662	CAATTCGTAACCAACCAAGGCGGTGTAAAGCAAGCAGATCTCTCTT					337613
693	asnserpro	la	g	l	as	709

709 pthreserleIleThrsrCysGlyAspClyAsnValValLysGlnGluG 726  
 337552 CCAAAAGCTGAGCTAG.....GAGCAAAAAGTTAATGAACAGAA 337519

726 InLeuSerProLysLysLysGluAsnAsnAlaLeuLeuArgTyrLeuLeu 742  
::: :::::|||||:::|||||

33/316 CMICIGCACGAAGACAAAGAGAGAGACIGCT... 33/48/

337486 .....GAACCTTCTGAAGAA...AATGAAGACAGAGTCCCTGAGT 337449

```

337448 GGACGGAGAGAA.....GTCC 337432
776 erSerGIngluLyAspProLysIleuThrGlu.....ThrSer 790
337431 AAGAAAGCTGAAAGAAAAAGAAAGCAATAAAGAAAGCAAGACGCG 337382
791 GluGluGlySerGlyAspLeuAspAsnLeuAspAlaIleuGlyAspLeu 807
337381 GAAGAGCTGGAACAAGAAAGATATGCTGCTCCAGAAAGAAAGT 337332
807 urThrSerSerAspPheTyraSerAsnSerIleSerSerAsnGlySerIstL 824
337331 TACAGTGTGTA..... 337319
824 euGlyThrLySGInGlnValPheGlnGlyThrAsnSerLeuGlyLeuLys 840
337318 .....GAAAG 337313
841 SerSerGInSerValGInSerIleArgProProThyAsnArgAlaValSe 857
337312 GTTGAATTAGTGTGTTATTTCAGAGCTCCAGAAATCAAGCTAAATAC 337263
857 rLeuAspSerProValSerValGlySerSerProProValLysAsnIleS 874
337262 TGTACCTCAACCAACACACATCCCAACACACAGCA..... 337223
874 erAlaPheProMetLeuProLysGInProMetLeuGlyGlyAsnProArg 890
337222 .....CCACAGCAACACAGACAGCAACCAACAGCAACAGCAAG 337181
891 MetMet.AspSerGInGluAsnTyrglySerSerMetGlyGlyProAsnA 907
337180 CAACAACAACACACACACACACACACACACACACACACACACACAC 337131
907 rGAsn.ValThrValThrGlnThrProSerSerGlyAspTrpGlyLeuPr 923
337130 GCAACACACACACACACACACACACAGTACAGC..... 337098
923 oAsnSerLyAlaGlyArgMetGluProMetAsnSerAsnSerMetGlyA 940
337098 ..... 337098
940 rGProGlyGlyAspTyraThrSerLeuProArgProAlaLeuGlyGly 956
337097 .....TCAAGCTCAAGCCCAACAGACACATTAATCTCA 337065
957 .....SerIleProThrLeuProLeuArgSerAsnSerIlePr 969
337064 AACTACTATATCTCAACACAGCAACATATAGCTCAACACAGCATC 337015
969 ogLyAlaArgProVal.LeuGlnGlnGlnGlnMetLeuGlnMetArg 985
337014 AGTTACGACAACGATTTGTCCCAACACACATATGCTCAACACAG 336965
986 ProGlyGlyIleProMetGlyMetGlyAlaAsnProTyrglyGlnAlaI 1002
336964 CAA...CAGCATCCACAACCTCAATCAACAACCTCAATCAAGCAAG 336918
1002 aAlaSerAsnGlnLeuGly..... 1008
336917 TCCCAAAAGTCMAAAACAGGAGCAACGTGGCTGCCCAACAGTACTACA 336868
1009 .....SerTrpProAspGlyMetLeuSer 1016
336867 TGTATCAAAACCAATTCCTGGATATCTTATCA...GGATTTGTTGAT 336821
1017 MetGluGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSe 1033
336820 TCACACAGATACGTTACGGTCAACATATCAGCAACTGCTCAAAACAA 336771
1033 rLeuAspAspLeuValGlyProProSerAsnLeuGlyGlnSerAspG 1050
: ..... 1050

```

```

336770 C.....GCTCAAACTAGTGTAAAGCTAAC 336745
1050 LuArgAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAla 1066
336744 AATATATTTTCCAAACAGGTTATGCTCAACAGCGCGCAACACTGCTGCT 336695
1067 ThrGlyLeuGluGluIleAspArgAlaLeuGlyIleProGluLeuValAs 1083
336694 GCTATATTTGACTGTGCTGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 336645
1083 ngGlnGlyGlnAlaLeuGluProTyrglyAspAlaPheGlnGlyGlnLys 1100
336644 TCAGGCCCA.....CCTCAACA.....CAACAGCCATACG 336613
1100 laAlaValMetMetAspGlnLysAlaGlyLeuTyrglyGlnThrTyPro 1116
336612 GTGGCTCATTCATCCATTAACCTACCCCATTTTACCAACAGTCAATCCCA 336563
1117 AlaGlnGlyProProMetGlnGlyGlyPheHisLeuGlnGlnSerPr 1133
336562 ...TATGCTCAACCTCAATACGGT.....GTACCTGCTCAATATCC 336525
1133 O.....SerPheAsnSerMetMetAsnGlnMetAsnG 1144
336524 ATACCACTTACCAAGAAAGCAATTCACACTATTCACAAACCTCAAAAGCTC 336475
1144 lngGlnGlyAsnPheProLeuGlnGlnGlnGlnGlnGlnGlnGlnGln 1160
336474 AGGAACAGCAAGTCCAAATCAAGTGT.....GCCAG 336440
1161 ArgProArgThrAsnThrProLysGlnLeuArgMetGlnGlnGlnGln 1177
336439 CATTTCTAAGACTCTCAACAGAGCAATTCACACAGCAACAGCAACGCA 336390
1177 gLeuGlnGlyGlnGlnPheLeuAsnGlnSerArgGlnAlaLeuGluLeu 1194
336389 ACCTCAAGGTCACACC..... 336374
1194 ysmetGluAsnProThrAlaGlyGlyAlaAlaValMetArgProMetMet 1210
336373 .....CAACCTGAAGTT 336362
1211 GlnProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1227
336361 CAAATGCMAAATGSC..... 336347
1227 gGluLeuLeuSerHisHisPheArgGlnArgValAlaMetMetMetG 1244
336347 ..... 336347
1244 lngGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1260
336346 ...CAACCTGTTAACCCACACACAAATGCAAGTTCACAACTACTATCA 336299
1261 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1270
336298 TTCCACACACACACAGCAACAGCTGCTGCC 336269
seq_name: /cgn2_6/plodata/2/pna/US09_NEW_COMB.seq:US-09-335-032-12216
seq_documentation_block:
: Sequence 12216, Application US/09335032
: GENERAL INFORMATION:
: APPLICANT: Velculescu, Victor
: APPLICANT: Vogelstein, Bert
: APPLICANT: Kinzler, Kenneth
: TITLE OF INVENTION: Characterization of the Yeast
: TITLE OF INVENTION: Transcriptome
: FILE REFERENCE: 01107.78572
: CURRENT APPLICATION NUMBER: US/09/335,032
: CURRENT FILING DATE: 1999-06-16
: PRIOR APPLICATION NUMBER: US 60/035,917
: PRIOR FILING DATE: 1997-01-23

```



```

515886 .....CAACCTATTGAAAAACAGGAAACAGCAATGTTTGCAT 515925
624 euleuthrCysSerSerAspArgIHisSerSerLeuThrAsnSer 640
||| ||| .....
515926 TATCCACTCCCTCAACAAACGATTAAAGAAAAAGAAAGACCTACT 515975
641 ProLeuAspSerSerCysIysGluSerSerValSerValThrSerProse 657
||||| ||| .....
515976 GAATGGAAAAATCTTACTGACAGAGAA..... 516005
657 rGlyValSerSerSerThrSerGlyIysValSerSerThrSerAsnMetH 674
516005 ..... 516005
674 lsgIySerLeuLeuGlnIuIySHsArgIleLeuHISLysLeuLeuGln 690
||||| ||| .....
516006 .....GCACCTAGTAGAAGAAAAAGAAAAAG..... 516029
691 AsnGlyAsnSerProAlaGluValAlaLysIleThrAlaGlnAlaThrG1 707
||||| ||| .....
516030 .....GAAAGAGCTGAGAAAAATACTGAGCAAAAT.. 516059
707 yLysAspThrSerSerIleThrSerCysGlyAspGlyAsnValIalysG 724
516060 .....GAG 516063
724 lngIuGlnLeuSerProLysLysIysGluAsnAspAlaLeuLeuArgTyr 740
||||| ||| .....
516064 AGGAGAAATTCACATGATACAGACGATTTAAAT.....TTAAGA... 516104
741 LeuLeuAspArgAspAspProSerAspAlaLeuSer.....LysG1 754
||||| ||| .....
516105 .....CTGGCAGTGCCTAGACACTGGTTAAAAAGA 516136
754 uLeuGlnProGlnValGlu...GlyValAspAsnLysMetSerGlnCysT 770
||||| ||| .....
516137 TCTTAAATTGGAAAGTGAAGTGCAGTCCCAATAGAGCCGCTTGCGA 516186
770 hrSerSerThrIleProSerSerSerGlnIuIyAspProLysIleLys 786
||||| ||| .....
516187 CTTTCTCATCCACTTATGATATCTCCCTCCCAAGCAAAAGATCATCAT 516236
787 ThrGluThrSerGlnGlnIySerGlyAspLeuAspAsnLeuAspAlaI1 803
||||| ||| .....
516237 ACCGAGACA.....GCTATGTATAGCTAATAAACTACTGATAT 516277
803 eLeuGlyAspLeuThrSerSerAspPheTyrAsn.....AsnS 816
||||| ||| .....
516278 CAATGAATACGTGACCCACACAAAGATTCCTCCCGGAAAGTATTAAT 516327
816 erIleSerSerAsnGlySerHisLeuGlyThrIySglnGlnIValPheGln 832
||||| ||| .....
516328 CAAATTGCGAATAATGGCAAAATGAATACAAACGAGAGTCGATTTGAT 516377
833 GlYThrAsnSerLeuGlyLeuLysSerSerGlnSerValGlnSerIleAr 849
||||| ||| .....
516378 TTCAATACTACT.....AAAAGGCTTCCTTTGTTACATCCATA.. 516416
849 gProProIyAsnArgAlaValSerLeuAspSerProValSerValIys 866
516417 .....CCGG 516420
866 erSerProProValIyAsnLysSerIalPheProMetLeuProLysGln 882
||||| ||| .....
516421 AAAGAGATCCCAAAAGAAATGCTCAGATGCTACATTT.....AAGCA 516464
883 ProMetLeuGlyIyAsnProArgMetMetAspSerGlnIuAsnIyrg1 899
||||| ||| .....
516465 CGTGAATCCCATGCAATGCGAGAGACTAAGT..... 516497
899 ySerSerMetGlyGlyProAsnArgAsnValThrValThrGlnThrProS 916
||||| ||| .....

```

```

516498 .....AATGTCAACATATCCATGCAACGAGAAAAACA 516528
916 erSerGlyAspTyrGlyLeu.....ProAsnSerLysAlaGlyArgMet 930
||||| ||| .....
516529 TCAACAAATGATACCTTTCTATATATAAAAAAATATCGGAGTGATTTA 516578
931 GluProMetAsnSerAsnSerMetGlyIyArgProGlyGlyAspTyrAsnTh 947
||||| ||| .....
516579 TCACAGCTTCTCATATGCTCCAGCTCTCCGTAATGCTCATCTCCA 516628
947 rSerLeuProArg.....P 952
516629 TACCTGCCCTGGAATATGATTCGCCCTCAGAAATCGATATTTGATC 516678
952 roAlaLeuGlyGlySerIleProThrLeuProLeuArgSerAsnSerIle 968
||||| ||| .....
516679 CGCATATGATGATGTCA.....CCCTCCAAACAAATACAGCC 516716
969 ProGlyAlaArgPro..... 973
516717 CCAGTTAGCTACGATTAAGATGCGATGATGCTTCAGAGAAAGACTTCAG 516766
974 .....ValLeuGlnGlnGlnGlnMetLeuGlnMetArgProG 987
516767 TTTGATATATACTTTAGCCAAACCGTATGACCGTTATATCTAGAGAG 516816
987 yGluIleProMetGlyMetGlyAlaAsnProTyrGlyGlnAlaAla 1003
||||| ||| .....
516817 GAGACATTAAT.....TCCGCGGATCACT 516842
1004 SerAsnGlnLeuGlySerTyrProAspGlyMet...LeuSerMetGlnG1 1019
||||| ||| .....
516843 AGTGAGAGAAATAGCTCTCAGCCAAAGATGATATACATATATGTTGAGCA 516892
1019 nValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerLeuAspA 1036
||||| ||| .....
516893 ATGCAATTGATTAACAGAAACAAAGATTTAATG...AATGAGTTG... 516935
1036 splEuValGlyProProSerAsnLeuGlnGlnSerAspGluArgAla 1052
516935 ..... 516935
1053 LeuLeuAspGlnLeuHISThrLeuLeuSerAsnThrAspAlaThrGlyLe 1069
||||| ||| .....
516936 .....ACACTGCTGCTACAGAA..... 516953
1069 uGlnGluIleAspArgAlaLeuGlyIleProGluLeuValAsnGlnIyG 1086
||||| ||| .....
516954 .....TTAGCAGATTCGATTAAGAGGAAA 516978
1086 lAlaLeuGluProLysGlnAspAlaPheGlnGlnIyGlnIuAlaIalVal 1102
||||| ||| .....
516979 CTGAATTAGAAAGAAATACGCGCTGATGAAACCAATTAATTCAGCACA 517028
1103 MetMetAspGlnLysAlaGlyLeuTyrGlyGlnThrIyProAlaGlnG1 1119
||||| ||| .....
517029 TCTTTGACGATCTTCGAGGTGCTTTTCTGACTTTGAAAAGAA.. 517076
1119 yProProMetGlnGlnIyPheHisLeuGlnGlnIyGlnSerProSerPheA 1136
||||| ||| .....
517077 .....CTGAGAAAAAAATCCAGTAAATTG 517101
1136 snSerMetMetAsnGlnMetAsnGlnGlnIyAsnPheProLeuGlnGly 1152
||||| ||| .....
517102 TGCATTTGATCCACGCAATTAATGACAG.....AGATTAAAGAGA 517142
1153 MetHisProArgAlaAsnIleMetArgProArgThrAsnThrProLysG1 1169
||||| ||| .....
517143 TTCAATTGCGGAGAGCAATTATTAATCCAAAGAAAAATGACACC...AAAC 517189
1169 nLeuArgMetGlnLeuGlnIuArgLeuGlnIyGlnIuPheLeuAsnG 1186
||||| ||| .....
517190 AAGTTTATGGAATCTAGAGAAATGAAAACTTGAATTAATCTATTG 517239

```

1186 InSerArgGlnAlaLeuGluLeuLysMet 1195  
:::||||::: |||:::|  
517240 ATGAGAGAGATTCA...GAAATTGAAATG 517265

---





```

957 ACAAGGCAACAAGGTTACCAAGTCACTTTTCCACACAGCCAGC 1006
93 eutyrglygl...ThrTyProAlaInglyProPrometGlnlygly 108
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1007 AATGGCAACAAGGTCCTACCACTTACCACTTCCGAGCAGCAGCA 1056
109 PheHISLeuGlnGlyInSerProSerPheAnSerMetMetAsnGln.. 124
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1057 CAAGGCAACAAGGTTACCAAGTCACTTACGACAGCCAGCGG 1106
125 .....MetAsnGlnGlnlyAsnPhProLeu.... 133
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1107 ACAAGTGCACAGTGGGACAGGCAAGGCAAGGTTACCAATGGCA 1156
134 .....GlnGlyMetHis 137
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1157 CTTCGCCAGCAGCAGGACAGGCAACAGCTAGACCAAGGCAACA 1206
138 Pro 138
   |||
1207 CCA 1209

```

seq\_name: /cgn2\_6/prodata/1/lna/5b\_COMB.seq:US-08-061-376-4

```

seq_documentation_block:
; Sequence 4, Application US/08061376
; Patent No. 6175000
; GENERAL INFORMATION:
; APPLICANT: Evans, Glen A.
; APPLICANT: Djabbal, Malek
; APPLICANT: Sellerl, Licia
; APPLICANT: Parry, Pauline
; TITLE OF INVENTION: CHARACTERIZATION OF A CHROMOSOME 11Q23
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/061,376
; FILING DATE: 13-MAY-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter, Stephen E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: P41 9387
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)546-4737
; TELEFAX: (619)546-9392
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11907 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
; US-08-061-376-4

```

alignment\_scores:
Quality: 85.50 Length: 179
Ratio: 0.983 Gaps: 6
Percent Similarity: 48.603 Percent Identity: 21.788

alignment\_block:
US-09-041-994-2\_COPY\_1018\_1179 x US-08-061-376-4 ..
Align seg 1/1 to: US-08-061-376-4 from: 1 to: 11907

```

4 SerHis.....GlyThrGlnAsnArgProLeuLeuArgAsnSerLeuAs 18
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6829 AGTCACCTGGATGATCTTCATCTTCAGAAATGAAGACAGTCAGTTC 6878
18 PAspleuValGlyProProSerAsnLeuGlnGlyInSerAspGlnArgA 35
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6879 AGACTTGGGTCCAGAGCTCTCTTAAAGGACAGAAAGACCAAA.... 6924
35 lAleuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThrGly 51
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6925 .....GTCTGAGTTCCAAAGAGCTCAGAGGGA 6951
52 LeuGlnGlnIleAspArgAlaLeuGlyTleProGlnLeuValAsnGln.. 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6952 TCTGCACATATGATGCTTACCTGGAAATCTCTAAAGTGGCCACAGGT 7001
68 .....GlyGlnAlaL 71
   |||||
7002 TCATACACACATCTAGAGACTGATGTTAGTAAATCGGCTCTTGG 7051
71 euGlnProLysGlnAspAlaPheGlnGlyGlnAlaAlaValMetMet 87
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7052 CTGAACCCCTTCAGTGTCTTCTCTAAAGAGGCCCTC..... 7092
88 AspGlnLysAlaGlyLeuTyrglyGlnThrTyrgProAlaGlnlyProBr 104
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7093 .....TCCTTCCA..... 7101
104 oMetGlnGlyGlyPheHisLeuGlnGlyGln..... 114
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7102 .....CACCTCAATTCAGAGGCAAGGAAGAATGTCAGACCAAC 7141
115 .....SerProSerPheAsnSerMetMetAsnGlnMetAsnGln 127
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7142 ACACAGATTCTACCCCAATCAGCAAACTCTCTCAGATGAAGATCTGAA 7191
128 GlnGlyAsnPheProLeuGlnGlyMetHisProArgAlaAsnIleMetAr 144
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7192 GTCAAAACCTTGACCTATCTGTGAATGAGCAACACATCTCATTTATCA 7241
144 gProArgThrAsnThrProLysGlnLeuArgMetGln 156
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7242 CGAACATATGCGATCTAGTTCAGAGATAGAGACAG 7278

```

seq\_name: /cgn2\_6/prodata/1/lna/5a\_COMB.seq:US-08-320-559-1

```

seq_documentation_block:
; Sequence 1, Application US/08320559
; Patent No. 5631335
; GENERAL INFORMATION:
; APPLICANT: Croce, Carlo
; APPLICANT: Canaan, Eli
; TITLE OF INVENTION: Diagnostics, Therapeutics and Methods for
; TITLE OF INVENTION: Detection and Treatment of Acute Leukemias
; TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5631335rls
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
; OPERATING SYSTEM: PC-DOS

```



## alignment\_block:

US-09-041-994-2\_COPY\_1018\_1179 x US-08-327-392-1 ..

Align seg 1/1 to: US-08-327-392-1 from: 1 to: 14255

```

4 SerHis.....GlyThrGlnAsnArgProLeuLeuArgAsnSerLeuAs 18
|||||
6655 AGTCACCTTGATGGATCTTCATCTTCAGAAATGAAGACATCGAGTGCCTTC 6704
18 pAspleuValGlyProProSerAsnLeuGlnGlyInSerAspGluArg 35
|||||
6705 AGACTGTGGTGCCAGAGAGCTCCTTTAAAGGAGAGAAAGACCANA.... 6750
35 lAleuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThrGly 51
|||||
6751 .....GTGCTGAGTTTCCAGAGCTCCAGAGGGA 6777
52 LeuGlnGluIleAspArgAlaLeuGlyIleProGluLeuValAsnGln.. 67
|||||
6778 TCTGCACATTAATGTGGCTTACCCTGGAAATTCCTAAACTGCCCCACAGGT 6827
68 .....GlyGlnAlaL 71
6828 TCATAACACACATCTAGAGACTGATGTAGTAAATCGGCTCCTTGG 6877
71 euGluProLysGlnAspAlaPheGlnGlyGlnAlaAlaValMetMet 87
|||||
6878 CTGAACCCCTTCACGTGCTTCTCTTAAAGAGGCCCTC..... 6918
88 AspGlnLysAlaGlyLeuTyrglyGlnThrTyrgProAlaGlnGlyProPr 104
|||||
6919 .....TCCTTCCCA..... 6927
104 oMetGlnGlyGlyPheHisLeuGlnGlyGln..... 114
|||||
6928 .....CACCTCCATTTAGAGGGGCAAGAAGATGTCAGACCAAC 6967
115 .....SerProSerPheAsnSerMetMetAsnGlnMetAsnGln 127
|||||
6968 ACACAGATTCTTACCCCAATCAGCAAACTCCTCTCCAGATCAAGATCTGAA 7017
128 GlnGlyAsnPheProLeuGlnGlyMetHisProArgAlaAsnIleMetAr 144
|||||
7018 GTCAAAACCTTGAACTATCTGGAATGAGCAACAGATCATCATTAATCA 7067
144 gProArgThrAsnThrProLysGlnLeuArgMetGln 156
|||||
7068 CGAACATATGGGATCTAGTCCAGAGATAGAGACAG 7104

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-306-691B-55

seq_documentation_block:
; Sequence 55, Application US/08306691B
; Patent No. 5734039
; GENERAL INFORMATION:
; APPLICANT: Calabretta, Bruno
; APPLICANT: Skorski, Tomasz
; TITLE OF INVENTION: ANTISENSE
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavoorgna & Monaco, P.C.
; STREET: Two Penn Center, Suite 1800
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1

```

## CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: US/08/306,691B
; FILING DATE: September 15, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5734039e
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14255 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-306-691B-55

```

## alignment\_scores:

```

Quality: 85.50 Length: 179
Ratio: 0.983 Gaps: 6
Percent Similarity: 48.603 Percent Identity: 21.788

```

## alignment\_block:

US-09-041-994-2\_COPY\_1018\_1179 x US-08-306-691B-55 ..

Align seg 1/1 to: US-08-306-691B-55 from: 1 to: 14255

```

4 SerHis.....GlyThrGlnAsnArgProLeuLeuArgAsnSerLeuAs 18
|||||
6655 AGTCACCTTGATGGATCTTCATCTTCAGAAATGAAGACATCGAGTGCCTTC 6704
18 pAspleuValGlyProProSerAsnLeuGlnGlyInSerAspGluArg 35
|||||
6705 AGACTGTGGTGCCAGAGAGCTCCTTTAAAGGAGAGAAAGACCANA.... 6750
35 lAleuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThrGly 51
|||||
6751 .....GTGCTGAGTTTCCAGAGCTCCAGAGGGA 6777
52 LeuGlnGluIleAspArgAlaLeuGlyIleProGluLeuValAsnGln.. 67
|||||
6778 TCTGCACATTAATGTGGCTTACCCTGGAAATTCCTAAACTGCCCCACAGGT 6827
68 .....GlyGlnAlaL 71
6828 TCATAACACACATCTAGAGACTGATGTAGTAAATCGGCTCCTTGG 6877
71 euGluProLysGlnAspAlaPheGlnGlyGlnAlaAlaValMetMet 87
|||||
6878 CTGAACCCCTTCACGTGCTTCTCTTAAAGAGGCCCTC..... 6918
88 AspGlnLysAlaGlyLeuTyrglyGlnThrTyrgProAlaGlnGlyProPr 104
|||||
6919 .....TCCTTCCCA..... 6927
104 oMetGlnGlyGlyPheHisLeuGlnGlyGln..... 114
|||||
6928 .....CACCTCCATTTAGAGGGGCAAGAAGATGTCAGACCAAC 6967
115 .....SerProSerPheAsnSerMetMetAsnGlnMetAsnGln 127
|||||
6968 ACACAGATTCTTACCCCAATCAGCAAACTCCTCTCCAGATCAAGATCTGAA 7017
128 GlnGlyAsnPheProLeuGlnGlyMetHisProArgAlaAsnIleMetAr 144
|||||
7018 GTCAAAACCTTGAACTATCTGGAATGAGCAACAGATCATCATTAATCA 7067

```



```
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04496
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Deluca Esq., Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: TJU-1242
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1425
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
ANTI-SENSE: NO
PCT-US94-04496-1
```

```
alignment_scores:
Quality: 85.50 Length: 179
Ratio: 0.983 Gaps: 6
Percent Similarity: 48.603 Percent Identity: 21.788
```

alignment\_block:

US-09-041-994-2\_COPY\_1018\_1179 x PCT-US94-04496-1 ..

Align seg 1/1 to: PCT-US94-04496-1 from: 1 to: 14255

```
4 SerHis.....GlyThrGlnAsnArgProLeuLeuArgAsnSerLeuAs 18
|||||
6655 AGTCACCTGGATGGATCTTCATCTTCAGAAATGAGACAGTCACGCTTC 6704
18 PASpleuValGlyProPserAsnLeuGlnGlyInSerAspGluAAsG 35
|||||
6705 AGACTGGGTGTCAGACGCTCTTTAAAGGAGAGAACCAAA.... 6750
35 lAlLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThrGly 51
|||||
6751 .....GTGCTGAGTTCACAGAGCTCAGAGGGA 6777
52 LeuGlnGluIleAspArgAlaLeuGlyIleProGluLeuValAsnGln.. 67
|||||
6778 TCTGCACATAATGTGGCTTACCCTGGAATTCCTAAACTGGCCCAAGGT 6827
68 .....GlyGlnAlaLeu 71
6828 TCATACACACATCTAGAGACTGAATGTTAGTAAATCGCTCTCTTG 6877
71 euGluProLysGlnAspAlaPheGlnGlyGlnGluAlaValaValMetMet 87
|||||
6878 CTGAACCTCTTCAGTGTGCTTCTCTAAAGAGCCCTC..... 6918
88 AspGlnLysAlaGlyLeuThrGlyGlnThrThrProAlaGlnGlyProBr 104
|||||
6919 .....TCCTTCCA..... 6927
104 omeGlnGlyGlyPheHisLeuGlnGlyGln..... 114
|||||
6928 .....CACCTTCATTGAGAGGGCAAGAGATGTCAGACCAAC 6967
115 .....SerProSerPheAsnSerMetMetAsnGlnMetAsnGln 127
|||||
6968 ACACAGATTCTACCAATCAGCAAACTCCTCTCCAGATGAGATGCTGAA 7017
```

```
128 GlnGlyAsnPheProLeuGlnGlyMetHisProArgAlaAsnIleMetArg 144
|||||
7018 GTCAAAACCTTGAGCTATCTGATATGAGCAACAGATCATCATATCA 7067
144 gProArgThrAsnThrProLysGlnLeuArgMetGln 156
|||||
7068 CGAACATATGGGATCTACTTCAGAGATAGGAGACAG 7104
```

seq\_name: /cgn2\_6/prodata/1/lna/5A\_COMB.seq:us-08-227-536-1

seq\_documentation\_block:

```
; Sequence 1, Application US/08227536
; Patent No. 5638784
; GENERAL INFORMATION:
; APPLICANT: Ecken, Richard
; APPLICANT: Ewen, Mark
; APPLICANT: Livingston, David
; TITLE OF INVENTION: NUCLEIC ACID, ENCODING TRANSCRIPTION
; TITLE OF INVENTION: FACTOR P300 AND USES OF P300
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weingarten, Schurigin, Gagnebin & Hayes
; STREET: Ten Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02109
```

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/227,536

FILING DATE: 14-APR-1994

CLASSIFICATION: 436

ATTORNEY/AGENT INFORMATION:

NAME: Williams Ph.D., Kathleen A.

REGISTRATION NUMBER: 34,380

REFERENCE/DOCKET NUMBER: DFCI-308XX

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 542-2290

TELEFAX: (617) 451-0313

INFORMATION FOR SEQ. ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 9046 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

FEATURE:

NAME/KEY: CDS

LOCATION: 1200..8441

US-08-227-536-1

```
alignment_scores:
Quality: 85.00 Length: 187
Ratio: 1.062 Gaps: 7
Percent Similarity: 42.781 Percent Identity: 23.529
```

alignment\_block:

US-09-041-994-2\_COPY\_1018\_1179 x US-08-227-536-1 ..

Align seg 1/1 to: US-08-227-536-1 from: 1 to: 9046

```
6 GlyThrGlnAsnArgProLeuLeuArgAsnSerLeuAspAspLeuValGln 22
|||||
7329 GGCACGTGTCTCAACAGCGCTTACAAACCTTTGGCGACTCTCAGGTC 7378
```

```

22 yProPseTAsnIleuGlnGlyGlnSerAspGluArgAlaIleuAspG 39
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7379 TCCACCTGCTCCCTCCAGCAGCAACG.....GTGCTTAGTA 7416
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
39 InIeuIsthrIleuSerAsnThrAspAlaThrGlyLeuGluIle 55
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7417 TCCTTCAC..... 7424

56 AspArgAlaLeuGlyIleProGluIleValAsnGlnGlnAlaLeuG 72
   :: ||| :: ||| :: ||| :: ||| :: ||| :: ||| :: |||
7425 .....GCCAACCCCGACGTGGT..... 7442

72 uProLysGlnAspAlaPheIleGlnGlnAlaAlaValMetMetasp 89
   ||| :: ||| :: ||| :: ||| :: ||| :: ||| :: |||
7443 .....GTCGATTCATCAAC 7458

89 InLysAlaGlyLeuTyrglyGlnThrTyrgProAlaGlnGlyPro..... 103
   ||| :: ||| :: ||| :: ||| :: ||| :: ||| :: |||
7459 AGCGGCTGCCAGTATGCCAACTCTAATCCACAACCACTCCCTGGGCGAG 7508

104 .....Prom 105

7509 CCTGGCATGCCCGAGGCGAGCCAGGGCTACAGCCACTACATGCCAGG 7558

105 tGlnGlyGlyPheIleGlnGlnGlnSerProSerPheAsnSerMet. 121
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7559 TCAGCAGGGGGTCTCACTCCATCCAGCCAGATGCAGAACTGATCAATGC 7608

122 .....MetAsnGlnMetAsnGlnGln 128
   :: ||| :: ||| :: ||| :: ||| :: ||| :: ||| :: |||
7609 AGCGGCGCTTCAGAGGGCTGGCGCTGCCGCCAGCAGCAACACAGCA 7658

129 GlyAsnPheProLeuGlnGlyMetHisProArgAlaAsnIleMetArg 145
   :: ||| :: ||| :: ||| :: ||| :: ||| :: ||| :: |||
7659 CTCACACCCACCATGGAGGAGGAGAGCCCGACGCTCAGCAGATGAACAT 7708

145 oArgThrAsnThr...ProLysGlnLeuArg.....MetGlnLeuGln 159
   :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7709 GAACCAACACCACTGCTTCACAAATTCGAGACATCTTGAGACGACAGC 7758

159 InArgLeuGln 162
   || :: |||
7759 AATGATGCCA 7769

seq_name: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:PCT-US95-04682-1

seq_documentation_block:
  : Sequence 1, Application PC/TUS9504682
  : GENERAL INFORMATION:
  : APPLICANT:
  : TITLE OF INVENTION: NUCLEIC ACID ENCODING TRANSCRIPTION
  : TITLE OF INVENTION: FACTOR P300 AND USES OF P300
  : NUMBER OF SEQUENCES: 13
  : CORRESPONDENCE ADDRESS:
  : ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
  : STREET: Ten Post Office Square
  : CITY: Boston
  : STATE: MA
  : COUNTRY: US
  : ZIP: 02109
  : COMPUTER READABLE FORM:
  : MEDIUM TYPE: Floppy disk
  : COMPUTER: IBM PC compatible
  : OPERATING SYSTEM: PC-DOS/MS-DOS
  : SOFTWARE: PatentIn Release #1.0, Version #1.25
  : CURRENT APPLICATION DATA:
  : APPLICATION NUMBER: PCT/US95/04682
  : FILING DATE:
  : CLASSIFICATION:
  : PRIOR APPLICATION DATA:
  : APPLICATION NUMBER: US 08/227,536
  : FILING DATE: 14-April-1994

```

```

1 CLASSIFICATION:
2
3 ATTORNEY/AGENT INFORMATION:
4
5 NAME: Holliday C. Heine, Ph.D.
6
7 REGISTRATION NUMBER: 34,346
8
9 REFERENCE/DOCKET NUMBER: DECT-308X9999
10
11 TELECOMMUNICATION INFORMATION:
12
13 TELEPHONE: (617) 542-2290
14
15 TELEFAX: (617) 451-0313
16
17 INFORMATION FOR SEO ID: 1:
18
19 SEQUENCE CHARACTERISTICS:
20
21 LENGTH: 9046 base pairs
22
23 TYPE: nucleic acid
24
25 STRANDEDNESS: single
26
27 TOPOLOGY: linear
28
29 MOLECULE TYPE: cDNA
30
31 HYPOTHETICAL: NO
32
33 ANTI-SENSE: NO
34
35 FEATURE:
36
37 NAME/KEY: CDS
38
39 LOCATION: 1200..8441
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006

```

[illegible]

alignment\_block:

Align seg 1/1 to: PCT-US95-04682-1 from: 1 to: 9046

[illegible]





alignment\_scores:  
 Quality: 84.00 Length: 122  
 Ratio: 1.647 Gaps: 3  
 Percent Similarity: 41.803 Percent Identity: 27.049

alignment\_block:  
 US-09-041-994-2\_COPY\_1018\_1179 x US-08-361-806A-1 ..

Align seg 1/1 to: US-08-361-806A-1 from: 1 to: 1071

```

73 ProlyscInaspAlaphegInglyInglu...AlaAlayaImetMetas 88
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
595 CCAGAGCATGACACATTCAGAGCAGTCCTGCTGCTATGCGGACTGC 644
88 pGlnInlyAlaGlyLeuTyrglyInThrTyrrProAlaGlnlyPro... 103
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
645 TCAGCCCCCACCCTCATCTATGGGCCACACAGCCAGCTTATAGCTCTAGTC 694
103 ..... 103
695 AGCAGCTCAGAGCTCTTCGGCATTCCTCGCATGTCAGTACTACTATCTCAG 744
104 .....PrometGlnI 107
745 CCACAGCAGCAGCCCTATGCTGTGTCATGCCCATTTTCAGCCCACTCAGAC 794
107 yGlyPheHisLeuGlnlyInserProSerPheAsnSerMetMetasng 124
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
795 AGGTTTCCTCCAGCCTGTGTGTCCTGCTCTTCCAAAGCAGATGAGAC 844
124 lmetAsngInglyAsnPh.....ProleuGlnlyMet 136
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
845 ATGCTAACACAGCAGCTGCTTCGCCACATCATCTCTCTGCGCCCATG 894
137 HisProAlaAsnIleMetArgProArgThraSnrThrProlysgInle 153
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
895 CACCCCGAGGCTCTGCATCCAGCCCTGAGCTCTTCCGCCCACTG 944
153 uArgMetGlnleuGln 158
|:|||||:|||||:
945 CCCTGTGCAGATGCAG 960

```

seq\_name: /cgn2\_6/ptodata/1/lna/PCTUS.COMB.seq:PCT-US95-16806A-1

```

seq_documentation_block:
; Sequence 1, Application PC/TUS9516806A
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: E2 Binding Proteins
; NUMBER OF SEQUENCES: 21
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/16806A
; FILING DATE: December 22, 1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/361,806
; FILING DATE: 22-DEC-1994
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1071 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1071
; PCT-US95-16806A-1

```

alignment\_scores:  
 Quality: 84.00 Length: 122  
 Ratio: 1.647 Gaps: 3  
 Percent Similarity: 41.803 Percent Identity: 27.049

alignment\_block:  
 US-09-041-994-2\_COPY\_1018\_1179 x PCT-US95-16806A-1 ..

Align seg 1/1 to: PCT-US95-16806A-1 from: 1 to: 1071

```

73 ProlyscInaspAlaphegInglyInglu...AlaAlayaImetMetas 88
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
595 CCAGAGCATGACACATTCAGAGCAGTCCTGCTGCTATGCGGACTGC 644
88 pGlnInlyAlaGlyLeuTyrglyInThrTyrrProAlaGlnlyPro... 103
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
645 TCAGCCCCCACCCTCATCTATGGGCCACACAGCCAGCTTATAGCTCTAGTC 694
103 ..... 103
695 AGCAGCTCAGAGCTCTTCGGCATTCCTCGCATGTCAGTACTACTATCTCAG 744
104 .....PrometGlnI 107
745 CCACAGCAGCAGCCCTATGCTGTGTCATGCCCATTTTCAGCCCACTCAGAC 794
107 yGlyPheHisLeuGlnlyInserProSerPheAsnSerMetMetasng 124
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
795 AGGTTTCCTCCAGCCTGTGTGTCCTGCTCTTCCAAAGCAGATGAGAC 844
124 lmetAsngInglyAsnPh.....ProleuGlnlyMet 136
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
845 ATGCTAACACAGCAGCTGCTTCGCCACATCATCTCTCTGCGCCCATG 894
137 HisProAlaAsnIleMetArgProArgThraSnrThrProlysgInle 153
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
895 CACCCCGAGGCTCTGCATCCAGCCCTGAGCTCTTCCGCCCACTG 944
153 uArgMetGlnleuGln 158
|:|||||:|||||:
945 CCCTGTGCAGATGCAG 960

```

seq\_name: /cgn2\_6/ptodata/1/lna/5A.COMB.seq:US-08-198-446B-8

```

seq_documentation_block:
; Sequence 8, Application US/08198446B
; Patent No. 5674996
; GENERAL INFORMATION:
; APPLICANT: Hartwell, Leland H.
; APPLICANT: Weinert, Ted A.
; APPLICANT: Pilon, Sharon E.
; APPLICANT: Groudine, Mark T.
; TITLE OF INVENTION: Cell Cycle Checkpoint Genes
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC
; STREET: 1420 Fifth Ave., Suite 2800
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101-2347
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/198,446B
; FILING DATE: 18-FEB-1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:

```



Align seq 1/1 to reverse of: US-08-870-693-8 from: 1 to: 1896

```

3 ValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerLeuAspAs 19
  :::::::::::::::::::: ||| ::| ::| ::| ::| ::| ::|
1074 TTGTCCTCCAGTGGCCCGACGAGAGTCCAGCATGGCCACCTCTGCTGCTCA 1025
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
19 PleuVal.....GlyProSerAsnLeuGlnGlyGln 31
  ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
1024 CTCGGTGCAGAGTTCATCCACACCCCTCTACTGCTCTCATGAGAAAT 975
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
31 er.....AspGluArgAla 35
  ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
974 CCGTTCGCCACAGTCCAGATTCTCCCGAATCTCCCAACAGAGGGCT 925
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
36 LeuLeuSpGlnLeuHisThrLeuLeuSerAsnThrAspAlaThrGlyLe 52
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
924 GTTAGGCCCGCAACCTTAACCTTAGTCTCGAGC..... 889
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
52 uGluGluLeuAspArgAlaLeuGlyLeuProGluLeuValAsnGlnGly 69
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
888 .....CAGACAGTCCCTCTCTCTACTATAACACACACGAGGC 852
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
69 In..AlaLeuGluProLysGlnAspAlaPheGlnGlyGlnGluAlaAla 84
  ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
851 AGTATCTCGATATCTTACAGTAACTTAATCAATGATTAGTTAAACAAT 802
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
85 ValMetMetAspGlnLysAlaGlyLeuTyGlyGlnThr..... 97
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
801 ACAGGAGTAAATCAAAATTAGAGCTTACAAATTAATCTCAATGAATCA 752
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
98 .....TyPro...AlaGlnGlyProPrometGlnGlyLys 109
  ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
751 GTCCGTACCAAGATACCCCAATGCTGTAGATTCCTCAACAGCTGCTC 702
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
109 heHisLeuGlnGlyGlnSerProSerPheAsnSer...MetMetAsnGln 124
  ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
701 AAGGACTAATGACACACGACGACCTCCACCCAGTGGCTCACTTAACCAA 652
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
125 MetAsnGlnGlnGlyAsnPheProLeuGln..... 134
  ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
651 ATGAAACACAAACTATGATCTTACAGCCCTCAGGAACTTAATGCGCTC 602
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
135 .....GlyMetHisProArgAlaAsnIleMetArgProArgThrAsn 149
  ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
601 TTCACCTCCCATGTCACCATGAAAGCAATGAGTAATCCACAGGACCTC 552
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
149 hrProLysGlnLeuArg 154
  ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
551 CTCCTCCCAAGTCAGG 535
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|

```

seq\_name: /cgn2\_6/ptodata/1/lna/5B\_COMB.seq:US-08-525-742-5

```

seq_documentation_block:
; Sequence 5, Application US/08525742
; Patent No. 5871742
; GENERAL INFORMATION:
; APPLICANT: Saito, Shuji
; APPLICANT: Ohkawa, Setsuko
; APPLICANT: Seki, Sakiko
; APPLICANT: Onawa, Ikuroh
; APPLICANT: Funato, Hiroh
; APPLICANT: Iritani, Yoshikazu
; APPLICANT: Aoyama, Shigem
; APPLICANT: Takahashi, Kiyohito
; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
; ADDRESSEE: NAUGHTON
; STREET: 1725 K Street, Suite 1000

```

```

; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525/742
; FILING DATE: 25-SEP-1995
; CLASSIFICATION: 424
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 05-074139
; FILING DATE: 31-MAR-1993
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 05-245625
; FILING DATE: 30-SEP-1993
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP94/00541
; FILING DATE: 31-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Mclelland, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 950811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-659-2930
; TELEFAX: 202-8870357
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1945 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 97..1038
; US-08-525-742-5

```

```

alignment_scores:
Quality: 79.00 Length: 154
Ratio: 1.113 Gaps: 6
Percent Similarity: 46.104 Percent Identity: 24.026

```

alignment\_block:

US-09-041-994-2\_COPY\_1018\_1179 x US-08-525-742-5 ..

Align seq 1/1 to: US-08-525-742-5 from: 1 to: 1945

```

21 ValGlyProProSerAsnLeuGlnGlyGlnSerAspGluArgAlaLeu 37
  ||| ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
367 GTCACCTCCCACTCAACAAGCTGAAGTTAAGACTGAACAATTAATGGCAC 416
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
37 uAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThrGlyLeuGln 54
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
417 ACAATTAGTAACAACATGATGTAGTACAAATCAAGTCGAGGTACTGAMC 466
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
54 LuLeuAsp..... 56
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
467 AAGTTGAAGGTATTATTACCTCTGTGTCAACACCAAGGAATCGCT 516
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
57 .....ArgAlaLeuGlyLeuProGluLeuValAsnGlnGlyGlnAl 70
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
517 CCAAGCTCTTCCACCAATGGGAGTCTTAAGTTATAGTCCAAACAGC 566
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
70 aLeuGluProLys.....GlnAspAlaPheGlnGlyGln 82
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
567 TGGTCATCAACACACGACGACGCTCCGATGAATGCTCATCCAGGCTCAC 616
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
82 LuAlaValMetMetAspGlnLysAlaGlyLeuTyGlyGlnThrTy 98
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|

```

```
617 CACGTCCTCCAAATGGGTAGTCCTAAG.....TTATTAGGTCCAAACCAA 660
      |||   ...   |||   |||   ...
99  ProAlaGlnGlyProPheMetGlnGlyGly...PheHisLeuGlnGlyG1 114
      :::::   |||||   :::::   :::::   |||||
661 GGTGGTCATCCAAAGACCATGGCCAAATGCTCCACAAACCAACCAAGTCC 710
      |||
114 nSerProSerPheAsnSerMetMetAsnGlnMetAsnGlnGlnGlyAsnP 131
      |||
711 AAGACCA.....ATGAACCTCTCAAGGCAT. 735
      |||
131 heProLeuGlnGlyMetHisProArgAlaAsnIleMetArgProArgThr 147
      |||||   ...   |||||
736 .....CCTGTCCTGGACACAGCTGGCCACGACCT 765
      |||
148 AsnThrProLys 151
      |||   |||   ...
766 AACGGCCACAA 777
```



51 G1yleuGluc1u1leaspargalaleuGly1leProGluDeuValnsn1 67  
|||||  
3287 GGCGTGGAGAAATTCACAGAGCTTTGGGATTCCTGAACCTGTCAATCA 3336  
67 nG1yGlnAlaLeuGluProLySGlnAspAlaPheGlnGlyGlnAlaA 84  
|||||  
3337 GGGACAGGGATTAGAGCCAAACAGGATGCTTCCAAAGGCCAAGAGCAG 3386  
84 lAvalMetLeAspGlnLySAlaG1yleuTyrg1yGlnThrTyProAla 100  
|||||  
3387 CAGTAATGATGATGCAGAAAGCCAGGATTATATGACACAGACATACCA 3436  
101 GlnG1yProProwetGlnGlyPheHn1sleuGlnG1yGlnSerProse 117  
|||||  
3437 CAGGGCCCTCCATGCAGAGAGGCTTTCATCTTCAGAGCAATCACACATC 3486  
117 rPheAsnSerMetLeAsnGlnMetAsnGlnGlnGlnGlnPheProLeuG 134  
|||||  
3487 TTTTAACCTGATGATGAATCAGATGAACACAGCAAGCAATTTTCCCTCC 3536  
134 lng1yMetH1SProwAlaAsn1leMetArpProArGThrsnThrPro 150  
|||||  
3537 MAGGATGCACCCACAGCCACACATCATGAGACCCCGACAAACACCCCC 3586  
151 LysGlnLeuArGMetGlnLeuGlnGlnArGLeuGln 162  
|||||  
3587 AAGCACTTAGAATGCAGCTTCAGACAGGCTGCAG 3622  
seq\_name: /cgn2\_6/ptodata/1/pna/US095A\_COMB.seq:US-09-513-066-1

seq\_documentation\_block:  
: Sequence 1, Application US/09513066  
: GENERAL INFORMATION:  
: APPLICANT: Chen, J. Don  
: APPLICANT: Leo, Christopher  
: APPLICANT: Li, Hui  
: TITLE OF INVENTION: NOVEL USES FOR THE RAC3 TRANSCRIPTIONAL COACTIVATOR OF  
: FILE OF INVENTION: STEROID NUCLEAR RECEPTORS  
: FILE REFERENCE: UMG-026CP  
: CURRENT APPLICATION NUMBER: US/09/513\_066  
: CURRENT FILING DATE: 2000-02-24  
: PRIOR APPLICATION NUMBER: USSN 09/041,994  
: PRIOR FILING DATE: 1998-03-13  
: PRIOR APPLICATION NUMBER: USSN 60/073,674  
: PRIOR FILING DATE: 1998-02-04  
: NUMBER OF SEQ ID NOS: 37  
: SOFTWARE: PatentIn Ver. 2.0  
: SEQ ID NO 1  
: LENGTH: 4496  
: TYPE: DNA  
: ORGANISM: Homo sapiens  
: FEATURE:  
: NAME/KEY: CDS  
: LOCATION: (86)..(4330)  
US-09-513-066-1

Alignment\_scores:  
Quality: 845.00 Length: 162  
Ratio: 5.216 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

Alignment\_block:

US-09-041-994-2\_COPY\_1018\_1179 x US-09-513-066-1 ..

Align seg 1/1 to: US-09-513-066-1 from: 1 to: 4496

1 G1uGlnValSerH1sG1yThrg1naAsnArGProLeuLeuArGAsnSer1e 17  
|||||  
3137 GAACAGATTCTCATGCACTCAAAATAGGCCCTCTTAGAATTCCTT 3186  
17 unspAspLeuValG1yProProSerAsnLeuGlnG1yGlnSerAspGlnA 34

|||||  
3187 GGATGATCTTGTGGGCCACCTTCCAACTGGAAAGCCAGAGTACGAA 3236  
34 rGAlaLeuAspGln1n1nHn1sThrLeuSerAsnThrAspAlaThr 50  
|||||  
3237 GAGCATTTATTTGGACAGCTGCACACTCTTCTCAGCAACAGATGGGACA 3286  
51 G1yleuGluc1u1leaspargalaleuGly1leProGluDeuValnsn1 67  
|||||  
3287 GGCGTGGAGAAATTCACAGAGCTTTGGGATTCCTGAACCTGTCAATCA 3336  
67 nG1yGlnAlaLeuGluProLySGlnAspAlaPheGlnGlyGlnAlaA 84  
|||||  
3337 GGGACAGGGATTAGAGCCAAACAGGATGCTTCCAAAGGCCAAGAGCAG 3386  
84 lAvalMetLeAspGlnLySAlaG1yleuTyrg1yGlnThrTyProAla 100  
|||||  
3387 CAGTAATGATGATGCAGAAAGCCAGGATTATATGACACAGACATACCA 3436  
101 GlnG1yProProwetGlnGlyPheHn1sleuGlnG1yGlnSerProse 117  
|||||  
3437 CAGGGCCCTCCATGCAGAGAGGCTTTCATCTTCAGAGCAATCACACATC 3486  
117 rPheAsnSerMetLeAsnGlnMetAsnGlnGlnGlnGlnPheProLeuG 134  
|||||  
3487 TTTTAACCTGATGATGAATCAGATGAACACAGCAAGCAATTTTCCCTCC 3536  
134 lng1yMetH1SProwAlaAsn1leMetArpProArGThrsnThrPro 150  
|||||  
3537 MAGGATGCACCCACAGCCACACATCATGAGACCCCGACAAACACCCCC 3586  
151 LysGlnLeuArGMetGlnLeuGlnGlnArGLeuGln 162  
|||||  
3587 AAGCACTTAGAATGCAGCTTCAGACAGGCTGCAG 3622  
seq\_name: /cgn2\_6/ptodata/1/pna/US6006\_COMB.seq:US-60-068-511-1

seq\_documentation\_block:  
: Sequence 1, Application US/60068511  
: GENERAL INFORMATION:  
: APPLICANT: Suen, Chen-Shian  
: APPLICANT: Frell, Donald E.  
: APPLICANT: Lytle, Richard C.  
: TITLE OF INVENTION: Cloning and Expression of a Nuclear  
: TITLE OF INVENTION: Receptor Coactivator Proteins and Uses Thereof  
: NUMBER OF SEQUENCES: 3  
: CORRESPONDENCE ADDRESS:  
: ADDRESS: American Home Products Corporation  
: STREET: One Campus Drive  
: CITY: Parsippany  
: STATE: New Jersey  
: COUNTRY: USA  
: ZIP: 07054  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/60/068,511  
: FILING DATE:  
: CLASSIFICATION:  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Walsh, Andrea C.  
: REGISTRATION NUMBER: 34,988  
: REFERENCE/DOCKET NUMBER: 97243-00  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 973-683-2169  
: TELEFAX: 973-683-4117  
: INFORMATION FOR SEQ ID NO: 1:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 4789 base pairs  
: TYPE: nucleic acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 185..4750  
US-60-068-511-1

alignment\_scores:  
Quality: 845.00 Length: 162  
Ratio: 5.216 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-041-994-2\_COPY\_1018\_1179 x US-60-068-511-1 ..

Align seg 1/1 to: US-60-068-511-1 from: 1 to: 4789

```
1 GluGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerLe 17
|||||
3236 GAACAAGTTTCTCATGGCAGCTCAAAATAGGCTCTTCTTAGAATTCCT 3285
17 uAspAspLeuValGlyProProSerAsnLeuGluGlnSerAspGluA 34
|||||
3286 GGATGATCTTGTGGGCCACCTTCCAACTGGAAGGCCAGAGTGCAGAA 3335
34 rGAlaLeuLeuAspGlnLeuHisThrLeuSerAsnThrAspAlaThr 50
|||||
3336 GAGCATTTATGGACAGCTGACACTCTTCACCAACACAGATGCCACA 3385
51 GlyLeuGluGlnIleAspArgAlaLeuGlyIleProGluLeuValAsnG 67
|||||
3386 GGCTGGAAGAAATTGACAGAGCTTTGGGCATTCCTGAACCTGTCAATCA 3435
67 nGlyGlnAlaLeuGluProLysGlnAspAlaPheGlnGlyGlnIuaIaa 84
|||||
3436 GGGACAGCATTTAGAGCCCAAAACAGAGATGCTTCCAAAGGCCAAGAGAG 3485
84 IValMetMetAspGlnLysAlaGlyLeuTyrglyGlnThrTyProAla 100
|||||
3486 CAGTAATGATGATCAGAAAGCAGATTATATGACACACATATACCCACA 3535
101 GlnGlyProPrometGlnGlyPheHisLeuGlnGlyAsnProLeuG 117
|||||
3536 CAGGGGCTCCAAATGCAAGAGGCTTTTCATCTTACAGGACATATCACCATC 3585
117 rPheAsnSerMetLeuAsnGlnMetAsnGlnGlnIleAsnProLeuG 134
|||||
3586 TTTAACTCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3635
134 InGlyMetHisProArgAlaAsnIleMetArgProArgThrAsnThrPro 150
|||||
3636 AAGGAATGACACCCAGAGCCCAACATCATGAGACCCCGGACAAACACCCCC 3685
151 LysGlnLeuArgMetGlnLeuGlnGlnArgLeuGln 162
|||||
3686 AAGCAACTTAGAATGCAGCTTCACAGCAGAGCTGCAG 3721
```

seq\_name: /cgn2\_6/ptodata/1/pna/US094\_COMB.seq:US-09-440-612-1

seq\_documentation\_block:  
; Sequence 1, Application US/09440612  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Lex M. Cowsett  
; TITLE OF INVENTION: ANTISENSE MODULATION OF SRC-3 EXPRESSION  
; FILE REFERENCE: RTS-0042  
; CURRENT APPLICATION NUMBER: US/09/440,612

CURRENT FILING DATE: 1999-11-15  
; NUMBER OF SEQ ID NOS: 47  
; SEQ ID NO 1  
; LENGTH: 6754  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (184)..(4422)  
US-09-440-612-1

alignment\_scores:  
Quality: 845.00 Length: 162  
Ratio: 5.216 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-041-994-2\_COPY\_1018\_1179 x US-09-440-612-1 ..

Align seg 1/1 to: US-09-440-612-1 from: 1 to: 6754

```
1 GluGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerLe 17
|||||
3220 GAACAAGTTTCTCATGGCAGCTCAAAATAGGCTCTTCTTAGAATTCCT 3269
17 uAspAspLeuValGlyProProSerAsnLeuGluGlnSerAspGluA 34
|||||
3270 GGATGATCTTGTGGGCCACCTTCCAACTGGAAGGCCAGAGTGCAGAA 3319
34 rGAlaLeuLeuAspGlnLeuHisThrLeuSerAsnThrAspAlaThr 50
|||||
3320 GAGCATTTATGGACAGCTGACACTCTTCACCAACACAGATGCCACA 3369
51 GlyLeuGluGlnIleAspArgAlaLeuGlyIleProGluLeuValAsnG 67
|||||
3370 GGCTGGAAGAAATTGACAGAGCTTTGGGCATTCCTGAACCTGTCAATCA 3419
67 nGlyGlnAlaLeuGluProLysGlnAspAlaPheGlnGlyGlnIuaIaa 84
|||||
3420 GGGACAGCATTTAGAGCCCAAAACAGAGATGCTTCCAAAGGCCAAGAGAG 3469
84 IValMetMetAspGlnLysAlaGlyLeuTyrglyGlnThrTyProAla 100
|||||
3470 CAGTAATGATGATCAGAAAGCAGATTATATGACACACATATACCCACA 3519
101 GlnGlyProPrometGlnGlyPheHisLeuGlnGlyAsnProLeuG 117
|||||
3520 CAGGGGCTCCAAATGCAAGAGGCTTTTCATCTTACAGGACATATCACCATC 3569
117 rPheAsnSerMetLeuAsnGlnMetAsnGlnGlnIleAsnProLeuG 134
|||||
3570 TTTAACTCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3619
134 InGlyMetHisProArgAlaAsnIleMetArgProArgThrAsnThrPro 150
|||||
3620 AAGGAATGACACCCAGAGCCCAACATCATGAGACCCCGGACAAACACCCCC 3669
151 LysGlnLeuArgMetGlnLeuGlnGlnArgLeuGln 162
|||||
3670 AAGCAACTTAGAATGCAGCTTCACAGCAGAGCTGCAG 3705
```

seq\_name: /cgn2\_6/ptodata/1/pna/US094\_COMB.seq:US-09-440-612-3

seq\_documentation\_block:  
; Sequence 3, Application US/09440612A  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Lex M. Cowsett  
; TITLE OF INVENTION: ANTISENSE MODULATION OF SRC-3 EXPRESSION  
; FILE REFERENCE: RTS-0042  
; CURRENT APPLICATION NUMBER: US/09/440,612A  
; CURRENT FILING DATE: 1999-11-15

```

; NUMBER OF SEQ ID NOS: 49
; SEQ ID NO 3
; LENGTH: 6754
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (184)..(4422)
US-09-440-612-3

alignment_scores:
  Quality: 845.00      Length: 162
  Ratio: 5.216         Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-041-994-2_COPY_1018_1179 x US-09-440-612-3  ..
Align seg 1/1 to: US-09-440-612-3 from: 1 to: 6754

1  GUGLNVAlSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerle 17
|||||
3220 GAACAAGTTTCTCATGCGACCTCAAAATAGGCGCTTCTTAGGAATTCCT 3269
17  uaspaSpleuValGlyProProSerAsnLeuGluGlyGlnSerAspGluA 34
|||||
3270 GGATGATCTTGTGGGCCACCTTCCAACTGGAAAGCCAGAGTGCAGAA 3319
34  rglAlaLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThr 50
|||||
3320 GAGCATTTATGGACCGACGTCGACACTCTTCTCAGCAACACAGATGCCACA 3369
51  G1yLeuGluGluLeuAspArgAlaLeuGlyIleProGluLeuValAsnG1 67
|||||
3370 GGCCTGGAGAAATTGACAGAGCTTGGGCATTCCTGAACTGTCAATCA 3419
67  nGlyGlnAlaLeuGluProLySGlnAspAlaPheGlnGlyGlnGluAla 84
|||||
3420 GGGACAGGCAATTAGACCCCAACAGGATGCTTCCAAAGGCCAAGAGCAG 3469
84  lAlaMetMetAspGlnLysAlaGlyLeuTyrGlyGlnThrTyrProAla 100
|||||
3470 CAGTAAATGATGATCAGAAAGCGAGATTATATGACAGACATACCCAGCA 3519
101  GlnGlyProProwMetGlnGlyPheHisLeuGlnGlyGlnSerProse 117
|||||
3520 CAGGGCCCTCCAAATGCAAGAGGCTTTCATCTTCAGGACCAATCACCATC 3569
117  rPheAsnSerMetMetAsnGlnMetAsnGlnGlnGlnLysAsnPheProLeuG 134
|||||
3570 TTTTAACTTATGATGAAATCAGATGAACCAAGCAATTTTCCCTCC 3619
134  InGlyMetHisProArgAlaAsnIleMetArgProArgThrAsnThrPro 150
|||||
3620 AAGGAATGCAACCCAGACGCAACATCATGAGACCCCGGACAAACACCCCC 3669
151  LysGlnLeuArgMetGlnLeuGlnGlnArgLeuGln 162
|||||
3670 AAGCAACTTAGAATGACGCTTCAGCAGAGGCTGCAG 3705

seq_name: /cgn2_6/ptodate/1/pna/US091_COMB.seq:US-09-125-635-1

seq_documentation_block:
; Sequence 1, Application US/09125635
; GENERAL INFORMATION:
; APPLICANT: THE UNITED STATES OF AMERICA represented by THE SE
; TITLE OF INVENTION: AIB1, A novel steroid receptor co-activator
; FILE REFERENCE: 49944
; CURRENT APPLICATION NUMBER: US/09/125, 635
; CURRENT FILING DATE: 1998-08-21
; PRIOR APPLICATION NUMBER: 60/049,728
; PRIOR FILING DATE: 1997-06-17
```

```

; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 6835
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(4463)
US-09-125-635-1

alignment_scores:
  Quality: 845.00      Length: 162
  Ratio: 5.216         Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-041-994-2_COPY_1018_1179 x US-09-125-635-1  ..
Align seg 1/1 to: US-09-125-635-1 from: 1 to: 6835

1  GUGLNVAlSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerle 17
|||||
3252 GAACAAGTTTCTCATGCGACCTCAAAATAGGCGCTTCTTAGGAATTCCT 3301
17  uaspaSpleuValGlyProProSerAsnLeuGluGlyGlnSerAspGluA 34
|||||
3302 GGATGATCTTGTGGGCCACCTTCCAACTGGAAAGCCAGAGTGCAGAA 3351
34  rglAlaLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThr 50
|||||
3352 GAGCATTTATGGACCGACGTCGACACTCTTCTCAGCAACACAGATGCCACA 3401
51  G1yLeuGluGluLeuAspArgAlaLeuGlyIleProGluLeuValAsnG1 67
|||||
3402 GGCCTGGAGAAATTGACAGAGCTTGGGCATTCCTGAACTGTCAATCA 3451
67  nGlyGlnAlaLeuGluProLySGlnAspAlaPheGlnGlyGlnGluAla 84
|||||
3452 GGGACAGGCAATTAGACCCCAACAGGATGCTTCCAAAGGCCAAGAGCAG 3501
84  lAlaMetMetAspGlnLysAlaGlyLeuTyrGlyGlnThrTyrProAla 100
|||||
3502 CAGTAAATGATGATCAGAAAGCGAGATTATATGACAGACATACCCAGCA 3551
101  GlnGlyProProwMetGlnGlyPheHisLeuGlnGlyGlnSerProse 117
|||||
3552 CAGGGCCCTCCAAATGCAAGAGGCTTTCATCTTCAGGACCAATCACCATC 3601
117  rPheAsnSerMetMetAsnGlnMetAsnGlnGlnGlnLysAsnPheProLeuG 134
|||||
3602 TTTTAACTTATGATGAAATCAGATGAACCAAGCAATTTTCCCTCC 3651
134  InGlyMetHisProArgAlaAsnIleMetArgProArgThrAsnThrPro 150
|||||
3652 AAGGAATGCAACCCAGACGCAACATCATGAGACCCCGGACAAACACCCCC 3701
151  LysGlnLeuArgMetGlnLeuGlnGlnArgLeuGln 162
|||||
3702 AAGCAACTTAGAATGACGCTTCAGCAGAGGCTGCAG 3737

seq_name: /cgn2_6/ptodate/1/pna/US093_COMB.seq:US-09-359-922-3962

seq_documentation_block:
; Sequence 3962, Application US/09359922
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL CONFIGS OBTAINED FROM VARIOUS CDNA
; FILE REFERENCE: 20411-752CON1
; CURRENT APPLICATION NUMBER: US/09/359, 922
; CURRENT FILING DATE: 1998-07-22
```



EARLIER APPLICATION NUMBER: US 09/205,155  
EARLIER FILING DATE: 1998-12-03  
NUMBER OF SEQ ID NOS: 13203  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 3962  
LENGTH: 6855  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(6855)  
OTHER INFORMATION: n = A,T,C or G  
US-09-359-922-3962

alignment\_scores:  
Quality: 845.00 Length: 162  
Ratio: 5.216 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-041-994-2\_COPY\_1018\_1179 x US-09-359-922-3962 ..

Align seg 1/1 to: US-09-359-922-3962 from: 1 to: 6855

```
1  GUGUINALSERHISGLYTHRGINASNARGPROLEUATGASNSERLE 17
|||||
3302  GAGCATATCTTGTGGCCACCTTCCAACTGGAAGCCAGACTGCAGAA 3351
|||||
17  UASPSAPLEUVALGLYPROPOSERASNLEUGLUGLINSERASPGIUA 34
|||||
3302  GAGCATATCTTGTGGCCACCTTCCAACTGGAAGCCAGACTGCAGAA 3351
|||||
34  RGALALEUASPSGLINLEUHSRHLSEUASERASNTHRASPALATHR 50
|||||
3352  GAGCATATCTTGTGGCCACCTTCCAACTGGAAGCCAGACTGCAGAA 3401
|||||
51  GILYLEUGLUGLULEASPARGALALEUGLYLEPROGILULEUVALASGL 67
|||||
3402  GGCCTGGAGAAATTCAGACAGCTTGGGCACTTCTGACGACACAGATGC 3451
|||||
67  NGLYGINALEUGLUPROLYSGINASPALAPHEGLINGLINGUALA 84
|||||
3452  GGGACAGCATTTAGAGCCCAACAGATGCTTCCAAAGCCAGAGACAG 3501
|||||
84  LAVALMETETASPSGLINLYSALAAGLYLEURYGLINTHRTYRPROLA 100
|||||
3502  CAGTAATGATGATCAGAAAGCAGGATTAATGACACACATATCCACACA 3551
|||||
101  GINGLYPROPMETGLINGLYPHEHISLEUGLINGLYINSERPROSE 117
|||||
3552  CAGGGGCTCCCAATGCAAGAGGCTTTCATCTTCAGGACCAATCCACATC 3601
|||||
117  RPHEASNSERMETETASGLINMETASGLINGLINGLYASNPHLEUG 134
|||||
3602  TTTTAACCTATGATGATGATGATGATGATGATGATGATGATGATGATG 3651
|||||
134  INGLYMETHISPROARGALASNILEMETARGPROARGTHRASNTHRPRO 150
|||||
3652  AAGGAATGCACCCAGACGACCAACATCATGAGACCCGAGCAAAACACCC 3701
|||||
151  LYSGLINLEUARGMETGLINLEUGLINGLINGLYLEUGLIN 162
|||||
3702  AAGCACTTAGAATGACCTTCAGCAGAGGCTGCAG 3737
```

seq\_name: /cgn2\_6/ptodata/1/pna/US093\_COMB.seq:us-09-359-922-3962

## seq\_documentation\_block:

; Sequence 3962, Application US/09359922A  
; GENERAL INFORMATION:  
; APPLICANT: Leshkowitz, Dena  
; APPLICANT: Liu, Jin  
; TITLE OF INVENTION: NOVEL CONFIGS OBTAINED FROM VARIOUS CDNA

TITLE OF INVENTION: LIBRARIES  
FILE REFERENCE: 20411-752CON1  
CURRENT APPLICATION NUMBER: US/09/359,922A  
CURRENT FILING DATE: 1999-07-22  
EARLIER APPLICATION NUMBER: US 09/205,155  
EARLIER FILING DATE: 1998-12-03  
EARLIER APPLICATION NUMBER: US 09/034,341  
EARLIER FILING DATE: 1998-02-13  
NUMBER OF SEQ ID NOS: 13203  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 3962  
LENGTH: 6855  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(6855)  
OTHER INFORMATION: n = A,T,C or G  
US-09-359-922-3962

alignment\_scores:  
Quality: 845.00 Length: 162  
Ratio: 5.216 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-041-994-2\_COPY\_1018\_1179 x US-09-359-922-3962 ..

Align seg 1/1 to: US-09-359-922-3962 from: 1 to: 6855

```
1  GUGUINALSERHISGLYTHRGINASNARGPROLEUATGASNSERLE 17
|||||
3302  GAGCATATCTTGTGGCCACCTTCCAACTGGAAGCCAGACTGCAGAA 3351
|||||
17  UASPSAPLEUVALGLYPROPOSERASNLEUGLUGLINSERASPGIUA 34
|||||
3302  GAGCATATCTTGTGGCCACCTTCCAACTGGAAGCCAGACTGCAGAA 3351
|||||
34  RGALALEUASPSGLINLEUHSRHLSEUASERASNTHRASPALATHR 50
|||||
3352  GAGCATATCTTGTGGCCACCTTCCAACTGGAAGCCAGACTGCAGAA 3401
|||||
51  GILYLEUGLUGLULEASPARGALALEUGLYLEPROGILULEUVALASGL 67
|||||
3402  GGCCTGGAGAAATTCAGACAGCTTGGGCACTTCTGACGACACAGATGC 3451
|||||
67  NGLYGINALEUGLUPROLYSGINASPALAPHEGLINGLINGUALA 84
|||||
3452  GGGACAGCATTTAGAGCCCAACAGATGCTTCCAAAGCCAGAGACAG 3501
|||||
84  LAVALMETETASPSGLINLYSALAAGLYLEURYGLINTHRTYRPROLA 100
|||||
3502  CAGTAATGATGATCAGAAAGCAGGATTAATGACACACATATCCACACA 3551
|||||
101  GINGLYPROPMETGLINGLYPHEHISLEUGLINGLYINSERPROSE 117
|||||
3552  CAGGGGCTCCCAATGCAAGAGGCTTTCATCTTCAGGACCAATCCACATC 3601
|||||
117  RPHEASNSERMETETASGLINMETASGLINGLINGLYASNPHLEUG 134
|||||
3602  TTTTAACCTATGATGATGATGATGATGATGATGATGATGATGATGATG 3651
|||||
134  INGLYMETHISPROARGALASNILEMETARGPROARGTHRASNTHRPRO 150
|||||
3652  AAGGAATGCACCCAGACGACCAACATCATGAGACCCGAGCAAAACACCC 3701
|||||
151  LYSGLINLEUARGMETGLINLEUGLINGLINGLYLEUGLIN 162
|||||
3702  AAGCACTTAGAATGACCTTCAGCAGAGGCTGCAG 3737
```

seq\_name: /cgn2\_6/ptodata/1/pna/US6020\_COMB.seq:us-60-209-009-198

```
seq_documentation_block:
; Sequence 198, Application US/60209009
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary
; APPLICANT: Pearson, Cecelia I.
; TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: PA-0027 P
; CURRENT APPLICATION NUMBER: US/60/209, 009
; CURRENT FILING DATE: 2000-06-01
; NUMBER OF SEQ ID NOS: 501
; SOFTWARE: PERL Program
; SEQ ID NO 198
; LENGTH: 7116
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No: 1094199.1
; NAME/KEY: unsure
; LOCATION: 3941-3993, 5899-5939, 6951
; OTHER INFORMATION: a, t, c, g, or other
US-60-209-009-198
```

## alignment\_scores:

Quality:	845.00	Length:	162
Ratio:	5.216	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

## alignment\_block:

US-09-041-994-2\_COPY\_1018\_1179 x US-60-209-009-198 ..

Align seg 1/1 to: US-60-209-009-198 from: 1 to: 7116

```
1  GUGGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerle 17
|||||
3252 GAACAGTTTCTCATGCGACTCAAAATAGGCTCTTCTTAGAATCCCT 3301
17  uaspaSplEuValGlyProProSerAsnLeuGluGlyGlnSerAspGlu 34
|||||
3302 GGATGATCTTGTGGGCCACCTTCCAACTGGAAAGCCAGATGCCA 3351
34  rGAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThr 50
|||||
3352 GAGCATTATTGGACGACGCTGCACACTCTTCTCAGCAACAGATGCCA 3401
51  GtyleuGluGluIleAspArgAlaLeuGlyIleProGluLeuValAsnG 67
|||||
3402 GGCCTGGAAGAAATTGACAGAGCTTTGGGCATTCTGAACTTGTCAATCA 3451
67  nGlyGlnAlaLeuGluProLysGlnAspAlaPheGlnGlyGlnGluAla 84
|||||
3452 GGGACAGGCAATTAGAGCCCAACAGATGCTTCCAAAGGCCAAGAGC 3501
84  lavalMetLeuAspGlnLysAlaGlyLeuTyrGlyGlnThrTyrProAla 100
|||||
3502 CAGTAATGATGGATCAGAAAGCGAGATTATATGGACAGACATACCAGA 3551
101  GlnGlyProProMetGlnGlyGlyPheHisLeuGlnGlyGlnSerProse 117
|||||
3552 CAGGGGCTCCCAATGCAAGAGGCTTTCAATCTTCAGGACAATCACATC 3601
117  rPheAsnSerMetLeuAsnGlnMetAsnGlnGlnGlyAsnPheProLeuG 134
|||||
3602 TTTTAACCTATGATGAATCAGATGAACCAAGCAATTTTCTCTCC 3651
134  lnglyMetHisProArgAlaAsnIleMetArgProArgThrAsnThrPro 150
|||||
3652 AAGGAATGCACCCAGACCAACATCATAGAGACCCGAGCAAAACACCCC 3701
151  LysGlnLeuArgMetGlnLeuGlnGlnArgLeuGln 162
|||||
3702 AAGCAACTTAGAATGCAGCTTCAGCAGAGGCTGCAG 3737
```

seq.name: /cgn2/6/ptodata/1/pna/US6021\_COMB.seq:US-60-213-360-7991

```
seq_documentation_block:
; Sequence 7991, Application US/60213360
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeli
; APPLICANT: Diep, Dinh
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide Poly
; FILE REFERENCE: GX-0014 P
; CURRENT APPLICATION NUMBER: US/60/213, 360
; CURRENT FILING DATE: 2000-06-21
; NUMBER OF SEQ ID NOS: 8347
; SOFTWARE: PERL Program
; SEQ ID NO 7991
; LENGTH: 7116
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No: 1094199.1
; NAME/KEY: unsure
; LOCATION: 3941-3993, 5899-5939, 6951
; OTHER INFORMATION: a, t, c, g, or other
US-60-213-360-7991
```

## alignment\_scores:

Quality:	845.00	Length:	162
Ratio:	5.216	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

## alignment\_block:

US-09-041-994-2\_COPY\_1018\_1179 x US-60-213-360-7991 ..

Align seg 1/1 to: US-60-213-360-7991 from: 1 to: 7116

```
1  GUGGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerle 17
|||||
3252 GAACAGTTTCTCATGCGACTCAAAATAGGCTCTTCTTAGAATCCCT 3301
17  uaspaSplEuValGlyProProSerAsnLeuGluGlyGlnSerAspGlu 34
|||||
3302 GGATGATCTTGTGGGCCACCTTCCAACTGGAAAGCCAGATGCCA 3351
34  rGAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThr 50
|||||
3352 GAGCATTATTGGACGACGCTGCACACTCTTCTCAGCAACAGATGCCA 3401
51  GtyleuGluGluIleAspArgAlaLeuGlyIleProGluLeuValAsnG 67
|||||
3402 GGCCTGGAAGAAATTGACAGAGCTTTGGGCATTCTGAACTTGTCAATCA 3451
67  nGlyGlnAlaLeuGluProLysGlnAspAlaPheGlnGlyGlnGluAla 84
|||||
3452 GGGACAGGCAATTAGAGCCCAACAGATGCTTCCAAAGGCCAAGAGC 3501
84  lavalMetLeuAspGlnLysAlaGlyLeuTyrGlyGlnThrTyrProAla 100
|||||
3502 CAGTAATGATGGATCAGAAAGCGAGATTATATGGACAGACATACCAGA 3551
101  GlnGlyProProMetGlnGlyGlyPheHisLeuGlnGlyGlnSerProse 117
|||||
3552 CAGGGGCTCCCAATGCAAGAGGCTTTCAATCTTCAGGACAATCACATC 3601
117  rPheAsnSerMetLeuAsnGlnMetAsnGlnGlnGlyAsnPheProLeuG 134
|||||
3602 TTTTAACCTATGATGAATCAGATGAACCAAGCAATTTTCTCTCC 3651
134  lnglyMetHisProArgAlaAsnIleMetArgProArgThrAsnThrPro 150
```

```
|||||
3652 AAGGAATGCACCCAGAGCCCAACATCATGAGACCCCGGACAAACACCCCC 3701
151 LysGlnLeuArgMetGlnLeuGlnGlnArgLeuGln 162
|||||
3702 AAGCAACTTAGAATGACAGCTTCAGCAGAGGCTGCAG 3737

seq_name: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:US-09-652-123-8947

seq_documentation_block:
; Sequence 8947, Application US/09652123
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1186-001
; CURRENT APPLICATION NUMBER: US/09/652,123
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,135
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 9796
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8947
; LENGTH: 7888
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-123-8947
```

```
alignment_scores:
Quality: 845.00 Length: 162
Ratio: 5.216 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment\_block:

US-09-041-994-2\_COPY\_1018\_1179 x US-09-652-123-8947 ..

Align seg 1/1 to: US-09-652-123-8947 from: 1 to: 7888

```
1 GluGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerLe 17
|||||
3218 GACACAGTTCTTCATGCGACACTCAAAATAGGCCCTCTTCAGGAATTCCT 3267
17 uAspAspLeuValGlyProProSerAsnLeuGlyGlnSerSpGua 34
|||||
3268 GGATGATCTTGTGGCCACCTTCCAACTGGAAGGCCAGTAGCAGAA 3317
34 rGAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThr 50
|||||
3318 GAGCATTAATTGACACGCTGACACTCTTCTCAGACACAGATGCCACA 3367
51 GlyLeuGlnGluIleAspArgAlaLeuGlyIleProGluLeuValAsnG1 67
|||||
3368 GGCGTGGAGAAATTTGACAGAGCTTTGGCATTCTGAACTTGTCAATCA 3417
67 nGlyGlnAlaLeuGluIleProLysGlnAspAlaPheGlnGlyGlnGlnAla 84
|||||
3418 GGGACAGGCAATTAGAGCCCAACAGAGATGCTTCCAAAGGCCAAAGAGAG 3467
84 lAValMetLeuAspGlnLysAlaGlyLeuTyrglyGlnThrTyProAla 100
|||||
3468 CAGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3517
101 GlnGlyProProMetGlnGlyGlyPheHisLeuGlnGlnGlnSerProSe 117
|||||
3518 CAGGGGCTCCCAATGCAAGAGGCTTTCATCTTCAGGAGACATCACCATC 3567
117 rPheAsnSerMetLeuAsnGlnMetAsnGlnGlnGlyAsnPheProLeuG 134
|||||
3568 TTTTAACCTCTATGATGATGATGATGATGATGATGATGATGATGATGAT 3617
134 lnglyMetHisProArgAlaAsnIleMetArgProArgThrAsnThrPro 150
|||||
```

```
|||||
3618 AAGGAATGCACCCAGAGCCCAACATCATGAGACCCCGGACAAACACCCCC 3667
151 LysGlnLeuArgMetGlnLeuGlnGlnArgLeuGln 162
|||||
3668 AAGCAACTTAGAATGACAGCTTCAGCAGAGGCTGCAG 3703

seq_name: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:US-09-652-127-9831

seq_documentation_block:
; Sequence 9831, Application US/09652127
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1183-001
; CURRENT APPLICATION NUMBER: US/09/652,127
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,134
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 10475
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9831
; LENGTH: 7888
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-127-9831
```

```
alignment_scores:
Quality: 845.00 Length: 162
Ratio: 5.216 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment\_block:

US-09-041-994-2\_COPY\_1018\_1179 x US-09-652-127-9831 ..

Align seg 1/1 to: US-09-652-127-9831 from: 1 to: 7888

```
1 GluGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerLe 17
|||||
3218 GACACAGTTCTTCATGCGACACTCAAAATAGGCCCTCTTCAGGAATTCCT 3267
17 uAspAspLeuValGlyProProSerAsnLeuGlyGlnSerSpGua 34
|||||
3268 GGATGATCTTGTGGCCACCTTCCAACTGGAAGGCCAGTAGCAGAA 3317
34 rGAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThr 50
|||||
3318 GAGCATTAATTGACACGCTGACACTCTTCTCAGACACAGATGCCACA 3367
51 GlyLeuGlnGluIleAspArgAlaLeuGlyIleProGluLeuValAsnG1 67
|||||
3368 GGCGTGGAGAAATTTGACAGAGCTTTGGCATTCTGAACTTGTCAATCA 3417
67 nGlyGlnAlaLeuGluIleProLysGlnAspAlaPheGlnGlyGlnGlnAla 84
|||||
3418 GGGACAGGCAATTAGAGCCCAACAGAGATGCTTCCAAAGGCCAAAGAGAG 3467
84 lAValMetLeuAspGlnLysAlaGlyLeuTyrglyGlnThrTyProAla 100
|||||
3468 CAGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3517
101 GlnGlyProProMetGlnGlyGlyPheHisLeuGlnGlnGlnSerProSe 117
|||||
3518 CAGGGGCTCCCAATGCAAGAGGCTTTCATCTTCAGGAGACATCACCATC 3567
117 rPheAsnSerMetLeuAsnGlnMetAsnGlnGlnGlyAsnPheProLeuG 134
|||||
3568 TTTTAACCTCTATGATGATGATGATGATGATGATGATGATGATGATGAT 3617
134 lnglyMetHisProArgAlaAsnIleMetArgProArgThrAsnThrPro 150
|||||
3618 AAGGAATGCACCCAGAGCCCAACATCATGAGACCCCGGACAAACACCCCC 3667
```

151 LysGlnLeuArgMetGlnLeuGlnGlnArgLeuGln 162  
|||||  
3668 AAGCAACTTAGATGACGCTTCAGCAGAGGCTGCAG 3703

seq\_name: /cgn2\_6/ptodata/1/pna/US096C\_COMB.seq:US-09-652-355-11058

seq\_documentation\_block:

; Sequence 11058, Application US/09652355  
; GENERAL INFORMATION:  
; APPLICANT: Shyjan, Andrew W.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES  
; TITLE OF INVENTION: THEREFOR  
; FILE REFERENCE: 1600.1192-001  
; CURRENT APPLICATION NUMBER: US/09/652,355  
; CURRENT FILING DATE: 2000-08-30  
; PRIOR APPLICATION NUMBER: 60/151,136  
; PRIOR FILING DATE: 1999-08-30  
; NUMBER OF SEQ ID NOS: 11227  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11058  
; LENGTH: 7888  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-652-355-11058

alignment\_scores:  
Quality: 845.00 Length: 162  
Ratio: 5.216 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-041-994-2\_COPY\_1018\_1179 x US-09-652-355-11058 ..

Align seg 1/1 to: US-09-652-355-11058 from: 1 to: 7888

1 GluGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSer 17  
|||||  
3218 GAACAAGTTTCATGCGCACTCAMAATAGCGCTTCCTTAGGAAATCCCT 3267  
17 uaspaarpleuValGlyProProSerAsnLeuGlnGlnSerAspGlu 34  
|||||  
3268 GGAATGATCTTTGGGCCACCTTCACACTGGAAGGCCAGAGTGACGAAA 3317  
34 rgaLaleuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThr 50  
|||||  
3318 GAGCATTATTGGACAGCGTGCACACTCTTCGACGAACACAGATGCCACA 3367  
51 GlyLeuGlnGluIleAspArgAlaLeuGlyIleProGlnLeuValAsnGln 67  
|||||  
3368 GGCTTGGAAGAAATTCACAGAGCTTTGGCATTCCTGAACTTGTCATCA 3417  
67 nGlyGlnAlaLeuGluProLysGlnAspAlaPheGlnGlnGlnAla 84  
|||||  
3418 GGGACAGCGATTAGAGCCCAACAGAGATGCTTCCAGGCCAAGAAGCAG 3467  
84 lavalmetetaspGlnLysAlaGlyLeuTyrglyGlnThrTyrrProAla 100  
|||||  
3468 CAGTATATGATGATCAGAAGCAGAGATTTATGACAGACATACCCAGCA 3517  
101 GlnGlyProProMetGlnGlnGlyPheHisLeuGlnGlnGlnSerPro 117  
|||||  
3518 CAGGGGCTTCCAATGCAAGGAGGCTTTCTTCCTTCAGGGCAATTCACCATC 3567  
117 rPheAsnSerMetetaspGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 134  
|||||  
3568 TTTTAACCTATGATGATCAATCAGATGAACAGCAGCAATTTTCTCTCC 3617  
134 lnglymethisProArgAlaAsnIleMetArgProArgThrAsnThrPro 150  
|||||  
3618 AAGGAATGCACCCAGCAGCAACATCATGAGACCCCGGACAAACACCCCC 3667

151 LysGlnLeuArgMetGlnLeuGlnGlnArgLeuGln 162  
|||||  
3668 AAGCAACTTAGATGACGCTTCAGCAGAGGCTGCAG 3703

seq\_name: /cgn2\_6/ptodata/1/pna/US096D\_COMB.seq:US-09-663-693-1028

seq\_documentation\_block:

; Sequence 1028, Application US/09663693  
; GENERAL INFORMATION:  
; APPLICANT: Gearing, David P.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES  
; TITLE OF INVENTION: THEREFOR  
; FILE REFERENCE: 1600.1195-001  
; CURRENT APPLICATION NUMBER: US/09/663,693  
; CURRENT FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/154,986  
; PRIOR FILING DATE: 1999-09-21  
; NUMBER OF SEQ ID NOS: 1340  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1028  
; LENGTH: 7888  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-663-693-1028

alignment\_scores:  
Quality: 845.00 Length: 162  
Ratio: 5.216 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-041-994-2\_COPY\_1018\_1179 x US-09-663-693-1028 ..

Align seg 1/1 to: US-09-663-693-1028 from: 1 to: 7888

1 GluGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSer 17  
|||||  
3218 GAACAAGTTTCATGCGCACTCAMAATAGCGCTTCCTTAGGAAATCCCT 3267  
17 uaspaarpleuValGlyProProSerAsnLeuGlnGlnSerAspGlu 34  
|||||  
3268 GGAATGATCTTTGGGCCACCTTCACACTGGAAGGCCAGAGTGACGAAA 3317  
34 rgaLaleuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThr 50  
|||||  
3318 GAGCATTATTGGACAGCGTGCACACTCTTCGACGAACACAGATGCCACA 3367  
51 GlyLeuGlnGluIleAspArgAlaLeuGlyIleProGlnLeuValAsnGln 67  
|||||  
3368 GGCTTGGAAGAAATTCACAGAGCTTTGGCATTCCTGAACTTGTCATCA 3417  
67 nGlyGlnAlaLeuGluProLysGlnAspAlaPheGlnGlnGlnAla 84  
|||||  
3418 GGGACAGCGATTAGAGCCCAACAGAGATGCTTCCAGGCCAAGAAGCAG 3467  
84 lavalmetetaspGlnLysAlaGlyLeuTyrglyGlnThrTyrrProAla 100  
|||||  
3468 CAGTATATGATGATCAGAAGCAGAGATTTATGACAGACATACCCAGCA 3517  
101 GlnGlyProProMetGlnGlnGlyPheHisLeuGlnGlnGlnSerPro 117  
|||||  
3518 CAGGGGCTTCCAATGCAAGGAGGCTTTCTTCCTTCAGGGCAATTCACCATC 3567  
117 rPheAsnSerMetetaspGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 134  
|||||  
3568 TTTTAACCTATGATGATCAATCAGATGAACAGCAGCAATTTTCTCTCC 3617  
134 lnglymethisProArgAlaAsnIleMetArgProArgThrAsnThrPro 150  
|||||  
3618 AAGGAATGCACCCAGCAGCAACATCATGAGACCCCGGACAAACACCCCC 3667  
151 LysGlnLeuArgMetGlnLeuGlnGlnArgLeuGln 162

|||||  
3668 AAGCACTTAGAATGCAGCTTCAGCAGAGGCTGCAG 3703

seq\_name: /cgn2\_6/ptodata/1/pna/US096E\_COMB.seq:US-09-698-014-5135

seq\_documentation\_block:  
: Sequence 5135, Application US/09698014  
: GENERAL INFORMATION:  
: APPLICANT: Gearing, David P.  
: APPLICANT: Holtzman, Douglas A.  
: APPLICANT: Kingsbury, Gillian A.  
: TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES  
: TITLE OF INVENTION: THEREFOR  
: FILE REFERENCE: 1600, 2014-001  
: CURRENT APPLICATION NUMBER: US/09/698,014  
: CURRENT FILING DATE: 2000-10-27  
: PRIOR APPLICATION NUMBER: 60/162,363  
: PRIOR FILING DATE: 1999-10-29  
: NUMBER OF SEQ ID NOS: 6098  
: SOFTWARE: FastSeq for Windows Version 4.0  
: SEQ ID NO 5135  
: LENGTH: 7888  
: TYPE: DNA  
: ORGANISM: Homo sapiens  
US-09-698-014-5135

alignment\_scores:  
Quality: 845.00 Length: 162  
Ratio: 5.216 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-041-994-2\_COPY\_1018\_1179 x US-09-698-014-5135 ..

Align seg 1/1 to: US-09-698-014-5135 from: 1 to: 7888

1 GUGInValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerLe 17  
|||||  
3218 GAACAGATTTCATGATGCAGCTCAAAATAGGCGCTTCTTAGGAATTCCT 3267  
17 uAspAspLeuValGlyProProSerAsnLeuGluGlyGlnSerAspGlu 34  
|||||  
3268 GGATGATCTTGTGGCCACCTTCACACTGGAGGCCAGAGTGAAGAAA 3317  
34 tGAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThr 50  
|||||  
3318 GAGCATTATTGACACGACGCTGCACACTCTTCAGCAACACAGATGCCACA 3367  
51 G1yLeuGluGluIleAspArgAlaLeuGlyIleProGluLeuValAsnG1 67  
|||||  
3368 GGCCTGGAGAAATGACAGAGCTTGGGCATTGCTGAACCTGTCAATCA 3417  
67 nG1yGlnAlaLeuGluProLySGlnAspAlaPheGlnGlyGlnAla 84  
|||||  
3418 GGGACAGGATTAAGACCCCAACAGATGCTTCCAGGCCAAGACAGCAG 3467  
84 lavalMetMetAspGlnLySAlaGlyLeuTyrcylGlnThrTyrrProAla 100  
|||||  
3468 CAGTATGATGATGCAGAGCAGATTAATGACAGACATACCAGCA 3517  
101 G1nglyProPromerGlnGlyGlyPheHisLeuGlnGlyGlnSerProse 117  
|||||  
3518 CAGGGGCTTCATTAAGAGAGGCTTTCATCTTCAGGACAATCACCATC 3567  
117 rPheAsnSerMetMetAsnGlnMetAsnGlnGlnGlyAsnProLeuG 134  
|||||  
3568 TTTTAACCTATGATGAATCAGATGAACAGCAGGCAATTTTCTCTCC 3617  
134 lnglyMetHisProArgAlaAsnIleMetArgProArgThrAsnThrPro 150  
|||||  
3618 AAGGAATGCACCCAGCAGCCACATCATGAGACCCCGACAACACACCCC 3667

151 LysGlnLeuArgMetGlnLeuGlnGlnArgLeuGln 162  
|||||  
3668 AAGCACTTAGAATGCAGCTTCAGCAGAGGCTGCAG 3703

---

OM of: US-09-041-994-2\_COPY\_1018\_1179 to: Pending\_Patents\_NA\_New:\* out\_format : pfs  
Date: May 1, 2001 9:44 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

## Command line parameters:

-MODEL=framet.p2n.model -DEV=rlp  
-O=/cgn2\_6/ptodata/2/pna/us09\_NEW.COMB.seq:US-09-196-296B-1 + 845.00 1599.01 1.2e-81 47  
-DB=Pending\_Patents\_NA\_New -QEXT=fastap -SUFFIX=p2n.rnpn  
-GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000  
-LOOPEXT=0.000 -GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000  
-XGAPEXT=0.500 -GAPOP=6.000 -GAPEXT=7.000 -XGAPOP=10.000  
-XGAPEXT=0.500 -DEEXT=6.000 -DEEXT=7.000 -START=1  
-MATRIX=blomsum62 -TRANS=human40.cai -List=45 -DOCALIGN=200  
-THRESHOLD=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=200000000  
-USER=US09041994\_@cgn1\_1175 -NCPU=6 -ICPU=3 -LONGLOG -NO\_XLPHY  
-WAIT -THREADS=1

## Search information block:

Query: US-09-041-994-2\_COPY\_1018\_1179  
Query length: 162  
Database: Pending\_Patents\_NA\_New:\*  
Database sequences: 201016  
Search time (sec): 114.180000

## Score list:

Sequence	Strid	Orig	ZScore	EScore	len	Documentation	
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-09-196-296B-1 +				845.00	1599.01	1.2e-81	47
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-09-540-212A-48616 +				123.50	236.17	9.6e-06	5
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-60-248-505-387 +				101.00	166.66	0.0716	5
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-60-248-505-322 +				101.00	150.34	0.5807	3
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-09-739-449-6007 +				92.00	156.33	0.2694	3
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-09-739-449-217 +				86.00	129.99	7.90	1
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-09-654-935A-99 +				84.00	90.42	1.3e+03	1
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-09-335-032-12214 +				81.00	136.36	3.49	31
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-09-801-833-6422 +				80.50	131.47	6.54	31
/cgn2_6/ptodata/2/pna/PCR_NEW.COMB.seq:PCR-US01-09226-10 +				79.00	130.12	7.77	2
/cgn2_6/ptodata/2/pna/PCR_NEW.COMB.seq:PCR-US01-08117-154 +				78.50	103.88	224.94	4
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-60-248-505-604 +				77.50	141.40	1.83	5
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-60-281-553-382 +				77.50	129.25	8.68	214
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-09-822-687-3 +				77.50	127.90	10.33	249
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-09-822-687-1 +				77.50	127.90	10.33	249
/cgn2_6/ptodata/2/pna/PCR_NEW.COMB.seq:PCR-US01-09226-10 +				77.00	135.36	3.97	972
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-09-823-059-2 +				77.00	133.30	5.17	122
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-09-739-449-8100 +				77.00	128.82	9.18	122
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-09-739-449-201 +				77.00	87.61	1.8e+03	2
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-09-783-514-2055 +				76.00	116.46	4.480	2
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-09-739-449-5127 +				75.50	136.44	3.46	3
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-09-739-449-214 +				75.50	80.95	4.2e+03	3
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-09-540-212A-15573 +				75.00	143.58	1.38	3
/cgn2_6/ptodata/2/pna/PCR_NEW.COMB.seq:PCR-US01-01339-6987 +				74.50	102.41	271.46	1
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-09-335-032-12216 +				74.50	69.21	1.9e+04	1
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-09-335-032-12219 +				74.00	68.03	2.2e+04	1
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-09-833-381-1079 +				73.50	134.30	4.54	166
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-09-663-561-2 +				73.50	123.79	17.50	166
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-09-828-769-143 +				73.00	133.62	5.64	5
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-09-336-313A-187 +				73.00	124.03	16.97	5
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-09-833-206-838 +				72.50	123.11	19.10	1
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-09-833-381-1904 +				72.50	122.73	20.05	1
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-09-833-381-1905 +				72.50	122.73	20.05	1
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-60-254-168-5 +				71.50	120.45	26.87	157
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-60-254-168-39 +				71.50	91.27	1.1e+03	41
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-09-739-449-6744 +				71.00	120.34	354.14	1
/cgn2_6/ptodata/2/pna/us09_NEW.COMB.seq:US-60-248-505-482 +				71.00	85.74	2.3e+03	68

/cgn2\_6/ptodata/2/pna/us09\_NEW.COMB.seq:US-60-248-505-500 - 71.00 85.74 2.3e+03  
/cgn2\_6/ptodata/2/pna/us09\_NEW.COMB.seq:US-60-248-505-84 + 71.00 83.26 3.2e+03  
/cgn2\_6/ptodata/2/pna/us09\_NEW.COMB.seq:US-09-739-449-218 - 71.00 67.87 2.2e+04  
/cgn2\_6/ptodata/2/pna/us09\_NEW.COMB.seq:US-60-281-553-192 + 70.50 110.81 92.41  
/cgn2\_6/ptodata/2/pna/us09\_NEW.COMB.seq:US-09-809-628-3 - 70.50 102.97 252.64

seq\_name: /cgn2\_6/ptodata/2/pna/us09\_NEW.COMB.seq:US-09-196-296B-1

## seq\_documentation\_block:

Sequence 1, Application US/09196296B  
GENERAL INFORMATION:  
APPLICANT: Suen, Chen-Shian  
APPLICANT: Frail, Donald E.  
APPLICANT: Lyttle, Richard C.  
TITLE OF INVENTION: Cloning and Expression of a Nuclear  
RECEPTOR COACTIVATOR PROTEIN AND USES THEREOF  
FILE REFERENCE: 0630/01376  
CURRENT FILING DATE: 1998-11-19  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 4789  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-09-196-296B-1

## alignment\_scores:

Quality: 845.00 Length: 162  
Ratio: 5.216 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-041-994-2\_COPY\_1018\_1179 x US-09-196-296B-1 ..

Align seg 1/1 to: US-09-196-296B-1 from: 1 to: 4789

1	GIUGLIVALSERHISGLYTHRGINASNATGPROLEULEUKYASNSERIE	17
3236	GAACCAAGTTCTTCATGACCTCAAAATAGCCCTTCTTGAAGTAATCCCT	3285
17	UASAPRLEUVALIGLYPROPROSERASNLEUGLUGINSERASNSGLA	34
3286	GGATGATCTTGTGGCCACCTTCCAACTGGAAGCCAGAGGAGGAGAA	3335
34	RGALALEULEUASRGINLEUHNISTHRLLEULEUSERASNTHASRALATGR	50
3336	GACCATTTATGACACAGCTGCACACTTCTTGCACACACAGATGCCACA	3385
51	GLYLEUGLUGLULEASPRARGALALEUGLYLEROGLEULEUVALANSGL	67
3386	GGCCTGGAAGAAATGACAGAGCTTGGGCACTTCTGCAACTTCTCATCA	3435
67	NGLYGINALALEUGLIPROLYSGINSLASPRALPRNEGINGLYGINLALAA	84
3436	GGGACAGGCAATGAGGCCCAACAGATGCTTCCAGGCAAGGCAAGGCAAG	3485
84	LAVALEUETASRGINLYSALAGLYLEUHYRGINLYGINTHYRPROALA	100
3486	CAGTATGATGATGATGACGAGGAGGATATATGACAGACATGCCACGCA	3535
101	GLINGLYPROBROMETGLINGLYLYPHENISLEUGLINGLYGINSERPROSE	117
3536	CAGGCGCTTCATGCAAGGAGGCTTTCATCTTCAAGGCAAAACACACATC	3585
117	IRPEASNSEMETMETASGLIMETASNGINGLYNGLYASNPHEROLENG	134
3586	TTTTAACTATGATGATGATGATGATGATGATGATGATGATGATGATGATG	3635
134	INGLYMETHISPRODALAASNILEMETARGPROARGTHIRASNTHIRPRO	150
3636	AAGGATGACACCCAGGAGGCAACATGATGATGATGATGATGATGATGATG	3685

151 LysGlnLeuArgMetGlnLeuGlnGlnArgLeuGln 162  
|||||  
3686 AAGCAACTTAGAATGACAGCTTCACAGCAGGCTGCAG 3721

seq\_name: /cgn2\_6/prodata/2/pna/US09\_NEW\_COMB.seq:US-09-540-212A-48616

seq\_documentation\_block:

; Sequence 48616, Application US/09540212A  
; GENERAL INFORMATION:  
; APPLICANT: Sellhauer, Jeffrey J.  
; APPLICANT: Delegeane, Angelo M.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Scuve, Laura L.  
; APPLICANT: Mullahy, Sara J.  
; APPLICANT: Naughton, Rebecca E.  
; TITLE OF INVENTION: POLYNUCLEOTIDES OF AIRWAY AND LUNG SYSTEM TISSUE  
; FILE REFERENCE: PD-1034 CIP  
; CURRENT APPLICATION NUMBER: US/09/540, 212A  
; NUMBER OF SEQ ID NOS: 67551  
; SOFTWARE: PERL Program  
; SEQ ID NO 48616  
; LENGTH: 274  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No: hu01311169  
US-09-540-212A-48616

alignment\_scores:

Quality:	123.50	Length:	81
Ratio:	2.245	Gaps:	3
Percent Similarity:	67.901	Percent Identity:	40.741

alignment\_block:

US-09-041-994-2\_COPY\_1018\_1179 x US-09-540-212A-48616 ..

Align seg 1/1 to: US-09-540-212A-48616 from: 1 to: 274

35 AAlaLeuLeuArgGlnLeuHisThrLeuLeuSerAsnThrAspAlaThrG1 51  
|||||  
27 GCTGTTCTTGACAGCTGGTATCCTCTCTAGTGGCAAGTTGAAACGGA 76  
51 YLeuGlnGluIleAspArgAlaLeuGlyIleProGluLeuValAsnGlnG 68  
||| |||  
77 GCTAGCTGAATAGACAGAGCTCTGGGAATTGACAAACTTGT...CAGG 123  
68 YGlnAlaLeuGluProLysGlnAspAlaPheGlnGlyGlnGlnAlaAla 84  
||| |||  
124 GGGGTGTAGATGTATTATCAGAGAGATTTCACACCAACAAGCAACG 173  
85 .....ValMetMetAspGlnLysAlaGlyLeuTyrglyGlnThrTyrr 99  
.....  
174 CCACCTTTGATCATGAGAGAAAGACCCACCTTATTCACGCTTACTC 223

99 cAlaGlnGlyPro.....PrometGlnGly 107  
::: |||  
224 TTCCTCTTCTCTACTGCCATCTCCCTAGCCCTTTCCCAAGGC 266

seq\_name: /cgn2\_6/prodata/2/pna/US60\_NEW\_COMB.seq:US-60-248-505-387

seq\_documentation\_block:

; Sequence 387, Application US/60248505  
; GENERAL INFORMATION:  
; APPLICANT: Beasley, Ellen  
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED  
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR  
; FILE REFERENCE: c1000918  
; CURRENT APPLICATION NUMBER: US/60/248, 505

; CURRENT FILING DATE: 2000-11-15

; NUMBER OF SEQ ID NOS: 1998

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 387

; LENGTH: 5132

; TYPE: DNA

; ORGANISM: human

US-60-248-505-387

alignment\_scores:

Quality:	101.00	Length:	141
Ratio:	1.347	Gaps:	9
Percent Similarity:	53.191	Percent Identity:	32.624

alignment\_block:

US-09-041-994-2\_COPY\_1018\_1179 x US-60-248-505-387 ..

Align seg 1/1 to: US-60-248-505-387 from: 1 to: 5132

20 LeuValGlyProProSerAsnLeuGlnGlyGlnSerAspGluArgAlaLe 36  
||| ||| |||  
3353 CTCGGCGGCGCACCGGAGACTGCAGCCGAGGTGT...GAGCGGAGGCT 3399  
36 uLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThrGlyLeuG 53  
||| ||| |||  
3400 ACTCCCGGCGCTGGGTACCTCGGCGACCTCCACCCCGACTCTCCACCG 3449  
53 YGlnIleAspArgAlaLeuGlyIleProGluLeuValAsn..... 66  
||| ||| |||  
3450 AG.....CGGCTTCGGGCTCGGCGCTGTGGCCAGGCGGCT 3490  
67 ..GlnGlyGlnAlaLeuGlu.....ProLysGlnAs 76  
|||||  
3491 GCCAAGGCGCAGGCTGTACCTGCTGCGTCTTGTGTCCCACTGAC 3540  
76 PAlaPheGlnGlyGln...GluAlaAlaValMetMetAspGlnLysAlaG 92  
.....  
3541 CCGCTCCCGCTCTCAGCCAGACGCTGCGCTGCGCACTGCGAGAAAGCG 3590  
92 Y.....LeuTyrglyGlnThrTyrrProAlaGlnGlyProPrometGln... 106  
||| |||  
3591 GGACATTTCTGGGTCGAGCTGTCTCCGAGCTCCCTCCACCGCAGCGGA 3640  
107 ..GlyIlePheHisLeuGlnGlyGlnSerProSerPheAsnSerMetMe 122  
|||||  
3641 CCGGCTGGCAGCTAT.....AGGAAACAGGCTTCGGAAGCTTGC 3681  
122 tAsnGlnMetAsnGlnGlnGlnGlyAsnPheProLeuGlnGlyMetHisPro 139  
|||||  
3682 CACCAGACAAACACG.....ACCCGAGAGAGCCGCCCTC 3716  
139 rgAlaAsnIleMetLarProArg 146  
3717 CTCGCCGCTTCCGCGAGCCCGCG 3739

seq\_name: /cgn2\_6/prodata/2/pna/US60\_NEW\_COMB.seq:US-60-248-505-222

seq\_documentation\_block:

; Sequence 222, Application US/60248505  
; GENERAL INFORMATION:  
; APPLICANT: Beasley, Ellen  
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED  
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR  
; FILE REFERENCE: c1000918  
; CURRENT APPLICATION NUMBER: US/60/248, 505  
; CURRENT FILING DATE: 2000-11-15  
; NUMBER OF SEQ ID NOS: 1998  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 222  
; LENGTH: 31818  
; TYPE: DNA



ORGANISM: human  
US-60-248-505-222

alignment\_scores:  
Quality: 101.00 Length: 141  
Ratio: 1.347 Gaps: 9  
Percent Similarity: 53.191 Percent Identity: 32.624

alignment\_block:

US-09-041-994-2\_COPY\_1018\_1179 x US-60-248-505-222 ..

Align seg 1/1 to: US-60-248-505-222 from: 1 to: 31818

```
20 LeuValGlyProProSerAsnLeuGlyGlnSerAspGluArgAlaLe 36
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
30039 CTCGGCGGCGCACCGAGGAACTGACCCGAGGTGT...GACGGGAGGT 30085
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
36 uLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThrGlyLeu 53
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
30086 ACTCCCGCGGCTGGTACGCTCCGGGCACTCCACCCCACTCCACCG 30135
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
53 LuGluIleAspArgAlaLeuGlyIleProGluLeuValAsn..... 66
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
30136 AG.....CGGGCTCTCGGCTCGCGCTCTGGCGACGCGGAGGCT 30176
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
67 ..GlnGlyGlnAlaLeuGlu.....ProLysGlnAs 76
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
30177 GCCCAAGGCCAGGCTCTACCTCTCCCTGGGCTCTTGTGCCCCCAGC 30226
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
76 palAPheGlnGlyGln..GluAlaAlaValMetMetAspGlnLysAla 92
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
30227 CCGCTCCCGCTCTACACCAAGACTGCGCTGGGCACTCGGAGAAAGCC 30276
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
92 Ly..LeuTyrGlyGlnThrTyrProAlaGlnGlyProPromeGln... 106
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
30277 GGCACAAATTCCTGACGCTCTCCGAGACTCCCTCTCCACCGCAGCA 30326
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
107 ..GlyGlyPheHisLeuGlnGlyGlnSerProSerPheAsnSerMet 122
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
30327 CCGGGTGGCAGCTAT.....AGGGAACAGGCTTCGGAAGTCTGCG 30367
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
122 LAsnGlnMetAsnGlnGlnLysAsnPheProLeuGlnGlyMetHisPro 139
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
30368 CAACCAAGACAAACAG.....ACCCGAGGAGGCCCTCTC 30402
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
139 rGluAlaAsnIleMetArgProArg 146
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
30403 CTCGCGCTTTCGCGAGCCCGG 30425
```

seq\_name: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-739-449-6007

seq\_documentation\_block:

Sequence 6007, Application US/09739449  
GENERAL INFORMATION:  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof  
FILE REFERENCE: 38-10(15490)C  
CURRENT APPLICATION NUMBER: US/09/739,449  
CURRENT FILING DATE: 2000-12-19  
PRIOR APPLICATION NUMBER: US 09/514,000  
PRIOR FILING DATE: 2000-02-23  
NUMBER OF SEQ ID NOS: 13351  
SEQ ID NO 6007  
LENGTH: 2349  
TYPE: DNA  
ORGANISM: Agrobacterium tumefaciens  
US-09-739-449-6007

alignment\_scores:  
Quality: 92.00 Length: 168

Ratio: 1.045 Gaps: 6  
Percent Similarity: 52.381 Percent Identity: 24.405

alignment\_block:

US-09-041-994-2\_COPY\_1018\_1179 x US-09-739-449-6007 ..

Align seg 1/1 to: US-09-739-449-6007 from: 1 to: 2349

```
23 ProProSerAsnLeuGlyGlnSer.....As 32
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1375 CCCCAGCCCAACATGAAACCAACCAAGCCAGACGTCGTCGCCAGCGCGCA 1424
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
32 pGluArgAlaLeuLeuAspGlnLeuHisThrLeu.....SerAsp 47
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1425 TCTGGAAACATGATGAAACCAACATCGAAGACTTGGCCGTTCCGGCAATC 1474
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
47 hrAspAlaThrGly.....LeuGluGluIleAspArgAlaLeuGlyIle 61
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1475 GTGACGCGGCGCAGAAATGCTCTCCGAGCTACAGCC..... 1512
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
62 ProGluLeuValAsnGlnGlyGlnAlaLeuGluProLysGlnAspAla 78
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1513 .....ATGATGAACAATCTGACAGGCTGCGCCCGCCAGCGCAGAGCCA 1556
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
78 eGlnGlyGlnGluAlaAlaValMetMetAspGlnLysAlaGlyLeuTyr 95
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1557 GCAGGACACCAACAGACGACCAAGATGCCGACAGATCGACACACTGG 1606
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
95 LysIleThrTyrProAlaGln.....GlyProPromeGlnGly 101
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1607 GCGAAATTCCTTACGACGACGACGACGAGTTGACGACGACTTCAGACT 1656
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
102 .....GlyProPromeGlnGly 108
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1657 GACCAAGCACTTCAGACCCGATGACGCGCGGACCCACAGCAGGCGG 1706
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
108 yPheHisLeuGlnGlyGlnSerProSerPheAsnSerMetMetAsnGln 125
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1707 CGAAGCGGAAGACGACGACGATGGCGAAGATGGCCACGCCACAGG 1756
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
125 eLAsnGlnGlnGlyAsnPheProLeuGlnGlyMetHisProAlaAsn 141
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1757 AAGGCCAGCAGGTCAACCGGTCAACCAAGCCAGATGTCTACAGAGCG 1806
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
142 IleMetArgProArgThrAsn...ThrProLysGlnLeuArgMetGlu 157
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1807 GGACAAACAGCGCAGACGACGATGACCGCGGAACAATTCGCGCAGCGCT 1856
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
157 uGln 158
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1857 GAAA 1860
```

seq\_name: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-739-449-217

seq\_documentation\_block:

Sequence 217, Application US/09739449  
GENERAL INFORMATION:  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof  
FILE REFERENCE: 38-10(15490)C  
CURRENT APPLICATION NUMBER: US/09/739,449  
CURRENT FILING DATE: 2000-12-19  
PRIOR APPLICATION NUMBER: US 09/514,000  
PRIOR FILING DATE: 2000-02-23  
NUMBER OF SEQ ID NOS: 13351  
SEQ ID NO 217  
LENGTH: 397803  
TYPE: DNA  
ORGANISM: Agrobacterium tumefaciens  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(397803)

```
; OTHER INFORMATION: unsure at all n locations
US-09-739-449-217

alignment_scores:
  Quality: 92.00      Length: 168
  Ratio: 1.045      Gaps: 6
  Percent Similarity: 52.381      Percent Identity: 24.405

alignment_block:
US-09-041-994-2_COPY_1018_1179 x US-09-739-449-217 ..
Align seg 1/1 to: US-09-739-449-217 from: 1 to: 397803

23 ProProSerAsnLeuGlnGlyGlnSer.....As 32
   |||.....:|||||:
31580 CCCCAGCCCAACATGACACCAACCAAGCCAGAACTGTCGCCGACGCGCA 31629
32 pGlnArgAlaLeuLeuAspGlnLeuHisThrLeu.....SerAsnT 47
   |...:.....:|||||:
31630 TCTGGAAACATGATGAACAGATGAGAACCTTCCCGCTCCGCAATC 31679
47 hrAspAlaThrGly.....LeuGlnGlnLeuAspArgAlaLeuGlyLe 61
   |||||:.....:|||||:
31680 GTGACGCGCGCCAGCAAACTGCTCTGCGACGTCACAGCC..... 31717
62 ProGlnLeuValAsnGlnGlyGlnAlaLeuGlnProGlyGlnAspAlaPh 78
   :|||:|||||:|||||:|||||:
31718 .....ATGATGAAACATCTGACAGCTGCGCGCCGCGCAAGCGCA 31761
78 eGlnGlyGlnGlnAlaAlaValMetMetAspGlnLeuGlyLeuTyrG 95
   |||||:|||||:|||||:|||||:
31762 CGAGGGACGACAAACAGACAGCAAGATCGCGCAGATCGACAAAGCTGG 31811
95 LysIleThrTyrProAlaGln..... 101
   |||:|||||:
31812 GCGAAATCTTCACAGCAGCAGAAAGTGTGACAGACATTTCAGCTC 31861
102 ..... 31911
31862 GACCAGCACTTCGAGCAGCAGCATGCGACGCCGACCCAGACAGCGCGG 31911
108 yPheHisLeuGlnGlnGlnSerProSerPheAsnSerMetMetAsnGlnM 125
   :|||:|||||:|||||:|||||:
31912 CGAAGCGCAACAGCAGCAGATGCGGAGATGCGCAGCGCCACAGG 31961
125 eCAsnGlnGlnGlyAsnProLeuGlnGlnGlnMetHisProArgAlaAsn 141
   :|||:|||||:|||||:|||||:
31962 AAGCGCACAAAGTCAACCGGCTCAACAAGCCAAATGTCACAGAGGGC 32011
142 IleMetArgProArgTyrAsn...ThrProLysGlnLeuArgMetGlnLe 157
   :|||:|||||:|||||:|||||:
32012 GGACAAACGCGGAGCGACAGATGACCGGAAATTTGCGCGAGCGCGCT 32061
157 uGln 158
   |||:
32062 GAAA 32065

seq_name: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-654-935A-99

seq_documentation_block:
; Sequence 99, Application US/09654935A
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ren, Felyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qiong A.
; APPLICANT: Xue, Aiding J.
; APPLICANT: Wehrman, Tom
```

```
; APPLICANT: Wang, Zhiwei
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 793
; CURRENT APPLICATION NUMBER: US/09/654,935A
; CURRENT FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 271
; SOFTWARE: pt_FL-genes Version 1.0
; SEQ ID NO 99
; LENGTH: 12279
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (8405)..(12226)
US-09-654-935A-99

alignment_scores:
  Quality: 86.00      Length: 126
  Ratio: 1.284      Gaps: 6
  Percent Similarity: 53.175      Percent Identity: 28.571

alignment_block:
US-09-041-994-2_COPY_1018_1179 x US-09-654-935A-99/rev ..
Align seg 1/1 to reverse of: US-09-654-935A-99 from: 1 to: 12279

2 GlnValSerHisGlyThrGlnAsnArgPro...LeuLeuArgAsnSerLe 17
   |||||:|||||:|||||:|||||:
11884 CAGGTGCCACAGGTGGCCAGCTGGCCCTGAACTCGACCAAGGTGAGCT 11935
17 uAspAspLeuValGlyProProSerAsnLeuGlnGlnGlnSerAspGlnA 34
   :|||:|||||:
11834 GCAGATCTGCTGGCGCAT..... 11816
34 rGlnAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThr...AspAla 49
   :|||:|||||:|||||:|||||:
11815 ..GTAGCCCTGTGATGACCTTCACGAAATGCTGATCTGCTTGAAACAG 11768
50 ThrGlyLeuGlnGlnIleAspArgAlaLeuGlyIleProGlnLeuValAs 66
   |||||:|||||:|||||:|||||:
11767 CTGACGCTGACGAACTGCACAGAGCC.....GGCATGTGGCTCA 11727
66 nGlnGlyGlnAlaLeuGlnProGlyGlnAspAlaPhe..... 78
   :|||:|||||:|||||:|||||:
11726 GCAGGCTGTGGCTTGAAGTGTGAAAGCAGACGCTCTTGAAGGCCACATC 11677
79 .....GlnGlyGlnGlnAlaAlaValMetMetAsp 88
   :|||:|||||:|||||:|||||:
11676 ATGACCTTCAGCTGCAGCAGCAGAGAGAAAGCCGCTGTACTTGTCTCAG 11627
89 GlnLysAlaGlyLeuTyrGlyGlnThrTyrProAlaGln...GlyProPr 104
   |||:|||||:|||||:|||||:
11626 GCAGCCCTGGGATGACAAATGTGAGAGGCCATGCTTGTATCCTGA 11577
104 oMetGlnGlyGlyPheHisLeuGlnGly 113
   :|||:|||||:|||||:|||||:
11576 GCTCCAGGACGCTCAGCACAATCCGGGCG 11549

seq_name: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-032-12214

seq_documentation_block:
; Sequence 12214, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: 01107,78572
; CURRENT APPLICATION NUMBER: US/09/335,032
```

; CURRENT FILING DATE: 1999-06-16  
 ; PRIOR APPLICATION NUMBER: US 60/035,917  
 ; PRIOR FILING DATE: 1997-01-23  
 ; PRIOR APPLICATION NUMBER: US 09/012,031  
 ; PRIOR FILING DATE: 1998-01-22  
 ; NUMBER OF SEQ ID NOS: 12219  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 12214  
 ; LENGTH: 666448  
 ; TYPE: DNA  
 ; ORGANISM: Saccharomyces cerevisiae  
 ; US-09-335-032-12214

alignment\_scores:  
 Quality: 84.00 Length: 177  
 Ratio: 0.884 Gaps: 8  
 Percent Similarity: 53.672 Percent Identity: 25.989

## alignment\_block:

US-09-041-994-2\_COPY\_1018\_1179 x US-09-335-032-12214 ..

Align seg 1/1 to: US-09-335-032-12214 from: 1 to: 666448

```

1 GluGlnValSerHisGlyThr..GlnAsnArgProLeuLeuArgAsnSe 16
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
365901 GAGCAGTGTGTTATACACCGCTCAGACAGTCTCCATACGAATAACA 365950
16 rLeuAspAspLeuValGlyProProSerAsnLeuGlnGlyIn..... 30
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
365951 TACTAATGATTCGAAATGCAAAACGAAAGTGTACCGCTCAGCAGCTG 366000
31 ..SerAspGluArgAlaLeuLeuAspGlnLeuHisThr.....LeuLeu 44
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
366001 CCTGTGAGTGGTGTATTGCTTATAGAACTTGTGTTACTA 366050
45 SerAsnThrAspAlaThrGlyLeuGlnGluLeuLeuAspArgAlaLeuGly 61
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
366051 GCTCTACTGTATGATCATATATGGGACACGACAAAGATCAAGTTCAT 366100
61 eProGluLeuValAsnGln.....GlyGln. 69
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
366101 ACCTTCTTGTGACGAATATCTCAATTTCTTTACTAGTGCGCAGC 366150
70 ..AlaLeuGlnProLysGlnAspAlaPheGlnGlnGlnAlaAlaVal 85
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
366151 CTCAGCTTCCCCCAGCTCAGCAACATCGCAGCCACCAAGCACCCACAA 366200
86 MetLeuAspGlnLysAlaGlyLeuTyGlyGlnThrTyProAlaGlnG 102
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
366201 AAGTTGCACAGACGACAGACCTATCTGTAATTTTCCACACA 366250
102 yProProMetGlnGlyPheHisLeuGlnGln.....SerPro 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
366251 GCAACCCGACACCCATCATATCATTTTGGCCAGTTCGCTGTAATG 366300
117 ePheAsnSerMetMetAsnGlnMetAsnGlnGlnGlnGlnGlnGlnGln 133
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
366301 GTTTTCATCT.....AGACAAAGATCAAGTCCGCCAGC 366332
134 GlnGlnMetHisProArgAlaAsnLeuMetArgProArgThrAsnThr.. 149
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
366333 GAGGCTATGTCTCCAGTGCACCTGCCATGTTTACTACACATCTACAAA 366382
150 ..ProLysGlnLeuArgMetGlnLeuGlnGln 159
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
366383 CCCCATAATGTGCACAGCAACACAAAGA 366413

```

seq\_name: /cgn2\_6/prcddata/2/pna/us09\_NEW\_COMB.seq.us-09-801-833-6422

seq\_documentation\_block:  
 ; Sequence 6422, Application US/09801833  
 ; GENERAL INFORMATION:

; APPLICANT: Glucksmann, M. Alexandra  
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A  
 ; FILE OF INVENTION: HUMAN BRAIN LIBRARY  
 ; FILE REFERENCE: 1600.1037-005  
 ; CURRENT APPLICATION NUMBER: US/09/801,833  
 ; CURRENT FILING DATE: 2001-03-13  
 ; PRIOR APPLICATION NUMBER: 09/371,168  
 ; PRIOR FILING DATE: 1999-08-10  
 ; PRIOR APPLICATION NUMBER: 60/095,907  
 ; PRIOR FILING DATE: 1998-08-10  
 ; PRIOR APPLICATION NUMBER: 60/103,145  
 ; PRIOR FILING DATE: 1998-10-05  
 ; NUMBER OF SEQ ID NOS: 8285  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 6422  
 ; LENGTH: 2055  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)..(2055)  
 ; OTHER INFORMATION: n = A,T,C or G  
 ; US-09-801-833-6422

alignment\_scores:  
 Quality: 81.00 Length: 157  
 Ratio: 1.038 Gaps: 5  
 Percent Similarity: 49.682 Percent Identity: 22.293

## alignment\_block:

US-09-041-994-2\_COPY\_1018\_1179 x US-09-801-833-6422/rev ..

Align seg 1/1 to reverse of: US-09-801-833-6422 from: 1 to: 2055

```

22 GlyProProSerAsnLeuGluGlyGlnSerAspGluArgAlaLeuAs 38
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1025 GCGCGGCCAGGCCCCCAGGAAGGACCCAGCAGCAGCAGTACAGTA 976
38 pGlnLeuHisThrLeuLeuSerAsnThrAspAlaThrGly..... 51
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
975 TGAGTGCATGACAAAGTCATGTGGCGATGGCAGAGGGAGGCCAGC 926
52 .....LeuGlnGluLeuAspArgAlaLeuGlyIleProGluLeuVal 65
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
925 GCCAGCCTGCGCTGCTGCTGCTGAGGGCCGCTGCGCAGTGC 876
66 AsnGlnGly.....GlnAlaLeuGluProLysGlnAspAlaPheGln 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
875 AACCGGCTGATGATCAGCAGCTGTGTATGATGACCTCTACAGAG 826
80 yGlnGlnAlaAlaValMetMetAspGlnLysAlaGlyLeuTyGlyGln 97
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
825 TCAGCAGCAGAGC.....CACAGTGAAG 803
97 hTyTyProAlaGlnGlyProProMetGlnGlyGlyPheHisLeuGln 113
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
802 ATCACCACGCCCCGGGCGCGCTGTTCCGGGAGAAAGTTGAGGCG 753
114 GlnSerProSerPheAsnSerMetMetAsnGlnMetAsnGln..... 127
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
752 AAGACGTCGAGCCCGCAGACACAGAAAGACAGATGGCGAAGAACCCCA 703
128 .....GlnGlyAsnPheProLeuGlnGlyMetHis..ProArg 139
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
702 AGAGGAAGCCCGAGAGGACAGAGTGAGAGTGGCTTCCACACAG 653
140 AlaAsnIleMetArgProArgThrAsnThrProLysGlnLeuArgMetG 156
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
652 GCAACACGAGCAGAGAGAGGCGGCTCCCGCAGAAAGAAATACCTG 603
156 nLeuGlnGlnArgLeuGln 162
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```



```
103 ProbroMetGlnGlyPheHisLeuGlnGlnSerProSerPheAs 119
1232 .....CAGGTACAGCCAGCAGGACATTTCACAGG 1260
119 nserMetMetAsnGlnMetAsnGlnGlnGlyAsnPheProLeuGlnGlyM 136
1261 CCCAGGACAGGTGCAGCTGCAGCAGCAGAGGCGAG...CCGCTGAAGCAGG 1307
136 ethisProArgAlaAsnIleMetArgProArgThrAsnThr.....Pro 150
1308 TGCAGCCACAG.....GTGCAGCCCGCCAGCAGCATTCACAGCCCCCA 1348
151 LysGlnLeuArgMetGlnLeuGlnGlnArgLeuGln 162
1349 AGCAGAGTGCGAGCTGCAGCTGCAGAGCAGCAGGTCAG 1384

seq_name: /cgn2_6/prodata/2/pna/US60_NEW_COMB.seq:US-60-248-505-604

seq_documentation_block:
; Sequence 604, Application US/60248505
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: c1000918
; CURRENT APPLICATION NUMBER: US/60/248,505
; CURRENT FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 1998
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 604
; LENGTH: 45333
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(45333)
; OTHER INFORMATION: n = A,T,C or G
US-60-248-505-604

alignment_scores:
Quality: 78.50 Length: 158
Ratio: 0.924 Gaps: 8
Percent Similarity: 53.797 Percent Identity: 25.949

alignment_block:
US-09-041-994-2_COPY_1018_1179 x US-60-248-505-604 ..
Align seg 1/1 to: US-60-248-505-604 from: 1 to: 45333

17 LeuAspAspLeuValGlyProProSerAsnLeuGlnGlnSerAspG1 33
||||| ||| ||||| ||||| ||| |||||
20767 TTACAGACTCTTTCAGCACCTACGAAAAATTAGAACAGCAGGTGAATCA 20816
33 uArgAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaT 50
||| ||||| ||||| ||||| |||||
20817 CAGC.....CAGCAGGAGCATAC.....AATCCCATGCG 20848
50 hArgLeuGlnGlnIleAspArgAlaLeuGlyIleProGlnLeu.Valas 66
||| ||||| ||||| ||||| |||||
20849 TGCCTTTAGCCAGATG.....AAAGTACCTCA 20877
66 nGlnGlyAlaLeuGlnProGlyGlnAspAlaPheGlnGlnGlnGly 82
||||| ||||| ||||| ||||| |||||
20878 GAAACACACACTGCTACAGCAGCAGCAGCAGCCAGCAGCAGCAGCAGC 20927
83 .....AlaAlaValMetMet 87
20928 AGCAGCCGGTTTACACCTTCAGCCCGCCAGCAGATATATGACACTTCAGCAG 20977
88 AspGlnLysAlaGlyLeuTyGlyGlnThrTyProAlaGlnGlyProPr 104
```

```
20978 CAGACAGCAGCAGCAGATCTCTACCAACCTTACCCCGCCAGCAG...CCGCC 21024
104 oMetGlnGlyPheHisLeuGlnGlnSerProSerPheAsnSerM 121
21025 GCATCCATTTCACAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 21074
121 etMetAsnGlnMetAsnGlnGlnGlnGlnAsnPheProLeuGlnGlnGlyMetHis 137
21075 AGTTTTCACAG.....CAACAGCTACAGTTTTCACAGCAACAGCAGTTCAT 21118
138 ProArgAlaAsnIleMetArgProArgThrAsn...ThrProGlnGln 153
21119 CCTCCACAGCAGCTGCATCGCCCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 21168
153 uArgMetGlnLeuGlnGlnArg 160
21169 GCAGCATGCCCTGCAGCAGCAG 21190

seq_name: /cgn2_6/prodata/2/pna/US60_NEW_COMB.seq:US-60-281-593-382

seq_documentation_block:
; Sequence 382, Application US/60281593
; GENERAL INFORMATION:
; APPLICANT: Amy Lasek
; TITLE OF INVENTION: GENES EXPRESSED IN LUNG CANCER
; FILE REFERENCE: PA-0045 P
; CURRENT APPLICATION NUMBER: US/60/281,593
; CURRENT FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 519
; SOFTWARE: PERL Program
; SEQ ID NO 382
; LENGTH: 551
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 383376.7
; NAME/KEY: unsure
; LOCATION: 317
; OTHER INFORMATION: a, t, c, g, or other
US-60-281-593-382

alignment_scores:
Quality: 77.50 Length: 105
Ratio: 1.314 Gaps: 6
Percent Similarity: 56.190 Percent Identity: 30.476

alignment_block:
US-09-041-994-2_COPY_1018_1179 x US-60-281-593-382/rev ..
Align seg 1/1 to reverse of: US-60-281-593-382 from: 1 to: 551

62 ProGlnLeuValAsnGln...GlyGlnAlaLeuGlnProGlyGlnAspAl 77
||||| ||||| ||||| ||||| |||||
364 CCAGAGTTAGTCAACAGCTGGGCGAGGCTGAGGCCAGCCAGGTCGCCN. 317
77 aPheGlnGlnGlnIleAlaAlaValMet.....MetAspG 89
||||| ||||| ||||| ||||| |||||
316 .....TGGGATGTGCTCTTCTCTCGAAGAGTGTTTCTAGAGC 277
89 InLysAlaGlyLeuTyGlyGlnThrTyProAlaGlnGlyProPro... 104
||||| ||||| ||||| ||||| |||||
276 TGAAGCCGCGCTTAACATGTAAATGCCACACAGAGGTAGGCGCCAGCAG 227
105 ...MetGlnGlyPheHisLeuGlnGlnGlnGlnAsnPheProLeuGlnGly 119
||||| ||||| ||||| ||||| |||||
226 GGGATGAGTGGGCTGCTCAAGTACAGCAGCCAGCAGTTCAGGCTCA 177
119 nserMetMetAsnGlnMetAsnGlnGlnGlnGlnAsnPheProLeuGlnGlyM 136
||||| ||||| ||||| ||||| |||||
176 AATCACCCTGACTGAGCAGCCAGCAGCAAGCAGCAGTTCGAGGAAGCAGGGA 127
```

```
136 echisProarg.....AlaasnlleMetarg 144
      |||||
126 CTCACCTCCCTCCTCAGCATGTGTTGAGCTGAGCTCAGATACATCCGG 77
145 ProargThrAsnThr 149
      |||||
76 CCAGAGAGCAGCAGCT 62

seq_name: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-822-687-3
```

## seq\_documentation\_block:

```
; Sequence 3, Application US/09822687
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria Alexandra
; TITLE OF INVENTION: 31939, A NOVEL HUMAN LEUCINE-RICH REPEAT
; FILE REFERENCE: 10448-038001
; CURRENT APPLICATION NUMBER: US/09/822,687
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,919
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2142
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-687-3
```

```
alignment_scores:
  Quality: 77.50      Length: 190
  Ratio: 0.881       Gaps: 9
  Percent Similarity: 46.316   Percent Identity: 24.737
```

```
alignment_block:
US-09-041-994-2_COPY_1018_1179 x US-09-822-687-3 ..
```

```
Align seg 1/1 to: US-09-822-687-3 from: 1 to: 2142
```

```
5 HtSGlYThrGlnAsnArgProLeuLeuArgAsnSerLeuAspLeuVa 21
||||| ||||| |||
1614 CACGGCACC.....CGCCCCGGCTCTCGCGGCCACGAGAGAGCGCT 1657
21 1.....GlyProPro 25
1658 TCACGGTCCCATCAGCATGTGACGAGAGACGCCCTCAAGGACCTGGAC 1707
25 eAsnLeuGlnGlnSerAspGluArgAlaLeuLeuAspGlnLeuHis 41
||||| ||||| |||||
1708 GACGTATGAGAGACCAAAATCATCATCGGCTGCTT...CGTGGCCAT 1754
42 ThrLeuSerAsnThrAspAlaThrGlyLeuGlnGluLeuAspArgAl 58
||||| ||||| |||||
1755 CACGTTTCATGGCGCGGTGATGCTCTGCGCTTCTTCAAGCTCGCAGAC 1804
58 aLeuGlyIleProGluLeuValAsnGlnGlnAlaLeuGlnProLysG 75
||||| ||||| |||||
1805 ACCACACAGCTCCACAA.....GCACCACGGGCCACGCGCACCGCTG 1845
75 lAspAlaIleheGlnGlnGlnGlnAlaAlaValMetMetAspGlnLysAla 91
||||| ||||| |||||
1846 GAGATCATCAACGTGAGGAGAGCGAGCCCGCGCTCGCGCTGCTCGCT 1895
92 GlyLeuThrGlyGln..... 96
|||||
1896 GGCCCGCGCGCGCGCGCTGGCCAGTGGGGGTGCTGGCGCGGAGACGCC 1945
97 .....ThyTrpProAlaGlnGlnGlnProPromeG 106
||||| ||||| |||||
1946 ACCTGGCCCTGCCCGCGCTGAGAGCAGACCACTCAACACACCACTAC 1995
```

```
106 lncGlyPheHisLeuGlnGlnSerProSerPheAsnSerMetMet 122
      |||
1996 GT...SGCTAGCGCTTTCAAGC.....GCATCAACAG 2024
123 AsnGlnMetAsnGlnGlnGlnAsnPheProLeuGlnGlyMetHisPro 139
      |||||
2025 CAGCAACCCCGAGCGGGGG.....CTCGGGGGGCAAGGCCCGCC 2065
139 gAlaAsnIleMetArgProArgThrAsnThrProLysGlnLeuArgMetG 156
      |||||
2066 CTGGCCTCAACTCCATCCACGAACTCTGCT..... 2096
156 lncGlnGlnArgLeuGln 162
2097 ..CTTCAAGAGCGCGCTCCAA 2114
```

```
seq_name: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-822-687-1
```

## seq\_documentation\_block:

```
; Sequence 1, Application US/09822687
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria Alexandra
; TITLE OF INVENTION: 31939, A NOVEL HUMAN LEUCINE-RICH REPEAT
; FILE REFERENCE: 10448-038001
; CURRENT APPLICATION NUMBER: US/09/822,687
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,919
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2493
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (187)...(2325)
US-09-822-687-1
```

```
alignment_scores:
  Quality: 77.50      Length: 190
  Ratio: 0.881       Gaps: 9
  Percent Similarity: 46.316   Percent Identity: 24.737
```

```
alignment_block:
US-09-041-994-2_COPY_1018_1179 x US-09-822-687-1 ..
```

```
Align seg 1/1 to: US-09-822-687-1 from: 1 to: 2493
```

```
5 HtSGlYThrGlnAsnArgProLeuLeuArgAsnSerLeuAspLeuVa 21
||||| ||||| |||
1800 CACGGCACC.....CGCCCCGGCTCTCGCGGCCACGAGAGAGCGCT 1843
21 1.....GlyProPro 25
1844 TCACGGTCCCATCAGCATGTGACGAGAGACGCCCTCAAGGACCTGGAC 1893
25 eAsnLeuGlnGlnSerAspGluArgAlaLeuLeuAspGlnLeuHis 41
||||| ||||| |||||
1894 GACGTATGAGAGACCAAAATCATCATCGGCTGCTT...CGTGGCCAT 1940
42 ThrLeuSerAsnThrAspAlaThrGlyLeuGlnGluLeuAspArgAl 58
||||| ||||| |||||
1941 CACGTTTCATGGCGCGGTGATGCTCTGCGCTTCTTCAAGCTCGCAGAC 1990
58 aLeuGlyIleProGluLeuValAsnGlnGlnAlaLeuGlnProLysG 75
||||| ||||| |||||
1991 ACCACACAGCTCCACAA.....GCACCACGGGCCACGCGCACCGCTG 2031
75 lAspAlaIleheGlnGlnGlnGlnAlaAlaValMetMetAspGlnLysAla 91
||||| ||||| |||||
```

```
2032 GAGATCATCAACGTGAGAGACGCTGCCCGCGCTCGGCCGTGTCGGT 2081
92 GlyLeuTyrGlyGln..... 96
||| |||
2082 GGGCGCGCGCGCGCGCGCTGGCCAGTGGGGGTGTGTGGCGGGGACAGCC 2131
97 .....ThyTyrProIaGlnGlyProPromeTg 106
||| ||| |||
2132 ACCTGGGCCCTGCCGCCCTGTGAGGAGAGACGCTCAACACACACACTAC 2181
106 InGlyGlyPheHisLeuGlnGlyInSerProSerPheAsnSerMetMet 122
:: ||| ::||| |||
2182 GT...GGCTGACGGCTTCACAGC.....GCACTACAG 2210
123 AsnGlnMetAsnGlnGlnGlyAsnPheProLeuGlnGlyMetHisProAr 139
::: ||| ::||| |||
2211 CAGCAACCCCGAGCGCGCGGGG.....CTGCGGGGGCAAGGCCCGCC 2251
139 gAlaAsnIleMetArProArGThrAsnThrProLysGlnLeuArGmetG 156
::: ||| ::||| |||
2252 CTGGCCCTCAACTCCATCCACGACACCTCTGCT..... 2282
156 InLeuGlnGlnArgLeuGln 162
||| ||| ||| ||| |||
2283 ..CTTCAAGAGCGGGCTCCAA 2300
```

**This Page Blank (uspto)**



OM of: US-09-041-994-2\_COPY\_613\_752 to: Issued\_Patents\_NA:\* out\_format : pfs  
Date: May 1, 2001 6:26 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

## Command line parameters:

```
-MODEL=frame+P2n.model -DEV=xlp  
-O/cg2_6/pdata/1/ina/5A.COMB.seq:US-07-667-276A-3 +  
-DB=Issued_Patents_NA -QWRT=fastap -SUFFIX=p2n.rml  
-GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000  
-LOOPEXT=0.000 -GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000  
-XGAPEXT=0.500 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000  
-YGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1  
-MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45 -DOALIGN=200  
-THR_SCORE=PCT -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTWT=Pfs -NORM=ext -MINLEN=0 -MAXLEN=200000000  
-USER=US09041994_ecg2_1_105 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY  
-WAIT -THREADS=1
```

## Search information block:

```
Query: US-09-041-994-2_COPY_613_752  
Query length: 140  
Database: Issued_Patents_NA:*  
Database sequences: 302621  
Search length: 87301344  
Search time (sec): 107.960000
```

## score\_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation	
/cg2_6/pdata/1/ina/5A.COMB.seq:US-07-667-276A-3 +				86.00	148.10	2017
/cg2_6/pdata/1/ina/5A.COMB.seq:US-08-209-747-1 -				83.50	138.88	2793
/cg2_6/pdata/1/ina/5A.COMB.seq:US-08-458-298-1 -				83.50	138.88	2793
/cg2_6/pdata/1/ina/5A.COMB.seq:US-08-320-559-29 +				82.00	133.42	3376
/cg2_6/pdata/1/ina/5A.COMB.seq:US-08-545-860D-29 +				82.00	133.42	3376
/cg2_6/pdata/1/ina/5A.COMB.seq:US-08-864-038A-1 -				81.00	130.90	3489
/cg2_6/pdata/1/ina/5A.COMB.seq:US-08-757-669A-20 +				81.00	104.27	32207
/cg2_6/pdata/1/ina/5A.COMB.seq:US-09-253-691-3 +				79.00	132.68	397
/cg2_6/pdata/1/ina/5A.COMB.seq:US-08-864-038A-2 -				78.50	131.03	2214
/cg2_6/pdata/1/ina/5A.COMB.seq:US-08-864-038A-1 -				78.50	126.14	3331
/cg2_6/pdata/1/ina/5A.COMB.seq:US-08-610-610-1 -				78.50	126.14	3331
/cg2_6/pdata/1/ina/5A.COMB.seq:US-08-719-822B-1 -				78.50	126.11	3337
/cg2_6/pdata/1/ina/5A.COMB.seq:US-08-820-170A-20 +				78.00	134.49	1518
/cg2_6/pdata/1/ina/5A.COMB.seq:US-09-055-699-20 +				78.00	134.49	1518
/cg2_6/pdata/1/ina/5A.COMB.seq:US-09-273-565-20 +				78.00	127.88	2636
/cg2_6/pdata/1/ina/5A.COMB.seq:US-08-820-170A-21 +				78.00	127.88	2636
/cg2_6/pdata/1/ina/5A.COMB.seq:US-09-055-699-21 +				78.00	127.88	2636
/cg2_6/pdata/1/ina/5A.COMB.seq:US-09-273-565-21 +				78.00	126.86	2870
/cg2_6/pdata/1/ina/5A.COMB.seq:US-08-468-036-28 +				78.00	126.86	2870
/cg2_6/pdata/1/ina/5A.COMB.seq:US-08-376-843-28 +				78.00	126.86	2870
/cg2_6/pdata/1/ina/5A.COMB.seq:US-07-853-913-1 +				78.00	110.51	11236
/cg2_6/pdata/1/ina/5A.COMB.seq:US-08-819-177-2 +				77.50	125.95	2832
/cg2_6/pdata/1/ina/5A.COMB.seq:US-08-776-265-6 +				77.50	129.11	1992
/cg2_6/pdata/1/ina/5A.COMB.seq:US-08-776-265-1 +				77.00	149.17	4565
/cg2_6/pdata/1/ina/5A.COMB.seq:US-09-197-649-7 -				76.00	146.52	390
/cg2_6/pdata/1/ina/5A.COMB.seq:US-09-10-928B-3 +				76.00	122.80	2824
/cg2_6/pdata/1/ina/5A.COMB.seq:US-08-469-802B-3 +				75.50	151.57	234
/cg2_6/pdata/1/ina/5A.COMB.seq:US-08-267-803B-3 +				75.50	151.57	234
/cg2_6/pdata/1/ina/5A.COMB.seq:US-08-209-747-7 -				75.50	135.84	870
/cg2_6/pdata/1/ina/5A.COMB.seq:US-08-458-298-7 -				75.50	127.54	1740
/cg2_6/pdata/1/ina/5A.COMB.seq:US-08-796-101-50 +				75.50	121.20	2953
/cg2_6/pdata/1/ina/5A.COMB.seq:US-08-859-201-1 +				75.50	113.85	5452
/cg2_6/pdata/1/ina/5A.COMB.seq:US-09-130-114-1 +				75.50	108.43	8575
/cg2_6/pdata/1/ina/5A.COMB.seq:US-07-884-811-15 -				75.50	105.89	10596
/cg2_6/pdata/1/ina/5A.COMB.seq:US-07-885-971-15 -				75.50	105.89	10596
/cg2_6/pdata/1/ina/5A.COMB.seq:US-08-087-783A-15 -				75.50	105.89	10596

```
seq_name: /cg2_6/pdata/1/ina/5A.COMB.seq:US-07-667-276A-3  
seq_documentation_block:  
; Sequence 3, Application US/07667276A  
; Patent No. 5470971  
; GENERAL INFORMATION:  
; APPLICANT: Kondo, Keiji  
; APPLICANT: Inouye, Masayori  
; TITLE OF INVENTION: STRESS-INDUCED PROTEINS, GENES CODING  
; TITLE OF INVENTION: THEREFOR, TRANSFORMED CELLS OF ORGANISMS, METHODS AND  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Weiser & Associates  
; STREET: 230 S. Fifteenth Street, Suite 500  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19102  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/667,276A  
; FILING DATE: 11-MAR-1991  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weiser, Gerard J.  
; REGISTRATION NUMBER: 19,763  
; REFERENCE/DOCKET NUMBER: 377,5551P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-875-8383  
; TELEFAX: 215-875-8394  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2017 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: Internal  
; ORIGINAL SOURCE:  
; ORGANISM: Saccharomyces cerevisiae  
; STRAIN: S288C  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 484..1725  
; NAME/KEY: misc_feature  
; LOCATION: 1  
; OTHER INFORMATION: /note="Base #1 of Sequence No. 5470971 3  
; OTHER INFORMATION: corresponds to base -483 of the sequence listed in  
; OTHER INFORMATION: Figure 11 of the application"  
; US-07-667-276A-3
```

## alignment\_scores:

Quality:	Length:
Ratio: 1.132	120
Percent Similarity: 63.333	Gaps: 2
Percent Identity: 24.167	

## alignment\_block:

```
US-09-041-994-2_COPY_613_752 x US-07-667-276A-3
```

Align seg 1/1 to: US-07-667-276A-3 from: 1 to: 2017

```

4 LysGlyHisLysLeuLeuGlnLeuLeuThrCysSerSerAspSpar 20
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
505 AAGGTAAACAGAGAGAGTAAAGCTTCCAAACAGCCAAAGAAAA 554
20 gclHisSerSerLeuThrAsnSerProLeuAspSerCysLysGln 37
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
555 AGCT...AAGGCCGCTCTCTCCCTCTCCGAACTCTCATCTCT 601
37 eSerSerSerValThrSerProSerGlyValSerSerThrSerGly 53
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
602 CATCTTCATCTGAATCTGAATCTGAGTCTGAGTCTGAATCTGA 651
54 GlyValSerSerThrSerAsnMetHisGlySerLeuLeuGlnLys 70
|||:|||||:|||||:|||||:|||||:|||||:|||||:
652 TCTTCATCTTCATCTCTGATCTGAACTCTCT... 604
70 sarGilleuHisLysLeuLeuGlnAsnGlyAsnSerProAlaGly 87
|||||:|||||:|||||:|||||:|||||:|||||:
685 .....TCTTCATCTCTCTCTGACAGCGAAAGTG 712
87 lalyslethrAlaGlnAlaThrGlyAspThrSerSerIleThrSer 103
|||:|||||:|||||:|||||:|||||:|||||:
713 AAGCTAAACCAAGAGAGAGAAATCCAGAGATCTCTCTCTCTCT 762
104 CysGlyAspGlyAsnValValLysGlnGlnGlnLeuSerProLys 120
|||||:|||||:|||||:|||||:|||||:|||||:
763 GACTCTTCTCCGACGAGAGAGAGAGAGAGAGAGAGAGAGAG 812
120 sGluAsnAsn 123
|||||:|||||:|||||:|||||:
813 GGAGAGATCA 822

```

seq\_name: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:US-08-209-747-1

```

seq_documentation_block:
: Sequence 1, Application US/08209747
: Patent No. 5733771
: GENERAL INFORMATION:
: APPLICANT: Lewis, Randolph V.
: TITLE OF INVENTION: CDNA Encoding Minor Ampullate Spider
: TITLE OF INVENTION: Silk Proteins
: NUMBER OF SEQUENCES: 56
: CORRESPONDENCE ADDRESS:
: STREET: P.O. Box 747
: CITY: Falls Church
: STATE: Virginia
: COUNTRY: USA
: ZIP: 22040-3487
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/209,747
: FILING DATE: 14-MAR-1994
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: Murphy Jr., Gerald M.
: REGISTRATION NUMBER: 28,977
: REFERENCE/DOCKET NUMBER: 1447-104P
: TELEPHONE: 703-205-8000
: TELEFAX: 703-205-8050
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2793 base pairs
: TYPE: nucleic acid

```

```

: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHETICAL: NO
: ORIGINAL SOURCE:
: ORGANISM: Nephila clavipes
: TISSUE TYPE: minor ampullate gland
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 183..2675
: OTHER INFORMATION: /product="N. clavipes minor
: OTHER INFORMATION: ampullate silk protein"
US-08-209-747-1

```

```

alignment_scores:
Quality: 83.50 Length: 116
Ratio: 1.176 Gaps: 4
Percent Similarity: 61.207 Percent Identity: 31.897

```

alignment\_block:

US-09-041-994-2\_copy\_613\_752 x US-08-209-747-1/rev ..

Align seg 1/1 to reverse of: US-08-209-747-1 from: 1 to: 2793

```

14 ThrCysSerSerAspSparGlyHisSerSer..... 24
|||||:|||||:|||||:|||||:|||||:|||||:
860 TCCTGTCTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 811
25 .....leuThrAsnSerProLeuAspSerSerCysLysGlnSer 39
|||||:|||||:|||||:|||||:|||||:|||||:
810 ACCCACCCTTGACCCGCTAGCCCTCCAGCTCCAGCAGCAGCAG 761
39 alSerValThrSerProSerGlyValSerSerSerThrSerGly 55
|||:|||||:|||||:|||||:|||||:|||||:
760 GCTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 711
56 SerSerThrSerAsnMetHisGlySerLeuLeuGln..... 67
|||||:|||||:|||||:|||||:|||||:|||||:
710 ACCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 662
68 .....GluLysHisArgLysLeuHisLysLeuLeuGln.Asn 79
|||:|||||:|||||:|||||:|||||:|||||:
661 TGTCACCATATTCACAGCAGCAGCAGCAGCAGCAGCAGCAG 612
80 GlyAsnSerProAlaGlnAlaLysIleThrAlaGlnAlaThrGly 96
|||||:|||||:|||||:|||||:|||||:|||||:
611 ACCAGCTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 562
96 sasPrThrSerSerIle.....ThrSerCysGlyAspGlyAsn 109
|||:|||||:|||||:|||||:|||||:|||||:
561 CTGCTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 516
seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-458-298-1
seq_documentation_block:
: Sequence 1, Application US/08458298
: Patent No. 5756677
: GENERAL INFORMATION:
: APPLICANT: Lewis, Randolph V.
: TITLE OF INVENTION: CDNA Encoding Minor Ampullate Spider
: TITLE OF INVENTION: Silk Proteins
: NUMBER OF SEQUENCES: 56
: CORRESPONDENCE ADDRESS:
: STREET: P.O. Box 747
: CITY: Falls Church
: STATE: Virginia
: COUNTRY: USA
: ZIP: 22040-3487
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk

```

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,298
FILING DATE: 02-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/209,747
FILING DATE: 14-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1447-104P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-205-8000
TELEFAX: 703-205-8050
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2793 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHEetical: NO
ORIGINAL SOURCE:
ORGANISM: Nephtila claviipes
TISSUE TYPE: minor amputate gland
FEATURE:
NAME/KEY: CDS
LOCATION: 183..2675
OTHER INFORMATION: /product= "N. claviipes minor
OTHER INFORMATION: amputate silk protein"

```

```
592 AGAAGTTGCTGAAGGCGAGGGGACCTTAATAGAGTATTCATTACCAG 641
16 rSerAspArgGlyHisSerSerLeuThrAsnSerProLeuAspSer 33
442 CAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 691
33 erCysLysGlnSerSerValSerValThrSerProSerGlyValSerSer 49
692 GCAGTACGACGACGACGACGACGACGACGACGACGACGACGACGACG 741
50 SerThrSerGlyGlyValSerSerThr..... 58
742 AGCAGCAGCAGCAGCAGCAGTAGTACCAAGTTTCCAAAGCCTCAAAATT 791
59 .....SerAsnMetHisGlys 64
792 AATGAGGACGACAGCAAGAAAACCTTCTAAAGACTCCAGAGAACATATAA 841
64 erLeuLeuGlnGlnLysHisArgGlyLeuHisLysLysLeuGlnAsnGly 80
842 GTGCCTTCAAGAAACCTTCCAGGGATCAACAACAAATCTTCCAAAGAAATCC 891
81 .....AsnSerProAlaGlnValAlaLysIleThrAl 91
892 TCTAAGAAACCCAAAGAAATATAACCTGAAAGAGAGAAATAGTTCC 941
91 aglnAlaThrGlyLysAspThrSer..... 99
942 TAAGATGCGCTCAAGAACCTTAACCACTGCAAAAGAGCCAAACACAG 991
100 .....SerIleThrSerCysGlyAspGlyAsn..... 108
992 ATACTAATTAATCTCAACCATCACGACAGACAGATTAAGAGGCTCTAGT 1041
109 .....ValValLysGlnGlnGlnLeuSerProLysLysGly 121
1042 AAAAGCCCCCATTTTCAGATTCTGAAGAACTCTCAGCCCAAAAAGGAA 1091
121 uAsnAsnAla 124
1092 AAAGAGTAGC 1101

seq_name: /cgn2_6/ptodata/1/lna/6A_COMB.seq:US-08-545-860D-29
seq_documentation_block:
; Sequence 29, Application US/08545860D
; Patent No. 6040140
; GENERAL INFORMATION:
; APPLICANT: Croce, Carlo
; APPLICANT: Canaan, Eli
; TITLE OF INVENTION: Diagnostics, Therapeutics and Methods
; TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias
; NUMBER OF INVENTION: Resulting from Chromosome Abnormalities in the All-1 Region
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
; ADDRESS: No. 6040140
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,860D
; FILING DATE: 07-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
```

```
APPLICATION NUMBER: PCT/US94/04496
FILING DATE: 22-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/10930
FILING DATE: 09-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/327,392
FILING DATE: 19-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/320,559
FILING DATE: 11-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/062,443
FILING DATE: 14-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/971,094
FILING DATE: 30-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/888,839
FILING DATE: 27-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/805,093
FILING DATE: 11-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Deluca Esq., Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: TJU-1262
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3439
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 3376 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 196..1902
US-08-545-860D-29

alignment_scores:
Quality: 82.00 Length: 170
Ratio: 1.000 Gaps: 5
Percent Similarity: 48.235 Percent Identity: 24.706

alignment_block:
US-09-041-994-2_COPY_613_752 x US-08-545-860D-29 ..
Align seq 1/1 to: US-08-545-860D-29 from: 1 to: 3376
7 LysLysLeuGlnGlnLeu.....LeuThrCys 16
16 rSerAspArgGlyHisSerSerLeuThrAsnSerProLeuAspSer 33
592 AGAAGTTGCTGAAGGCGAGGGGACCTTAATAGAGTATTCATTACCAG 641
442 CAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 691
33 erCysLysGlnSerSerValSerValThrSerProSerGlyValSerSer 49
692 GCAGTACGACGACGACGACGACGACGACGACGACGACGACGACGACG 741
50 SerThrSerGlyGlyValSerSerThr.....SerAsnMetHisGlys 64
742 AGCAGCAGCAGCAGCAGCAGTAGTACCAAGTTTCCAAAGCCTCAAAATT 791
59 .....SerAsnMetHisGlys 64
792 AATGAGGACGACAGCAAGAAAACCTTCTAAAGACTCCAGAGAACATATAA 841
```





```

      seq_name: /cgn2-6/pdata/1/lna/bB_COMB.seq:US-08-757-669A-20
seq_documentation_block:
  Sequence 20, Application US/08757669A
  Patent No. 6183751
  GENERAL INFORMATION:
    APPLICANT: Chang, Yvan
    APPLICANT: Bohenzky, Roy A.
    APPLICANT: Russo, James J.
    APPLICANT: Edelman, Isidore S.
    APPLICANT: Moore, Patrick S.
    TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
    TITLE OF INVENTION: SEQUENCES AND USES THEREOF
    NUMBER OF SEQUENCES: 20
    CORRESPONDENCE ADDRESS:
      ADDRESSEE: Cooper & Dunham LLP
      STREET: 1185 Avenue of the Americas
      CITY: New York
      STATE: New York
      COUNTRY: U.S.A.
      ZIP: 10036
    COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: Patentin Release #1.0, Version #1.30
      CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/08/757,669A
        FILING DATE:
          CLASSIFICATION: 424
        ATTORNEY/AGENT INFORMATION:
          NAME: White, John P.
          REGISTRATION NUMBER: 28,678
          REFERENCE/DOCKET NUMBER: 45185-F
          TELECOMMUNICATION INFORMATION:
            TELEPHONE: (212) 278-0400
            TELEFAX: (212) 391-0525
        INFORMATION FOR SEQ ID NO: 20:
          SEQUENCE CHARACTERISTICS:
            LENGTH: 32207 base pairs
            TYPE: nucleic acid
            STRANDEDNESS: double
            TOPOLOGY: linear
            MOLECULE TYPE: DNA (genomic)
            US-08-757-669A-20
alignment_scores:
  Quality: 81.00 Length: 108
  Ratio: 1.157 Gaps: 1
  Percent Similarity: 64.815 Percent Identity: 25.926
alignment_block:
  US-09-041-994-2_COPY_613_752 x US-08-757-669A-20 ..
Align seg 1/1 to: US-08-757-669A-20 from: 1 to: 32207
1 LeuGluSerLysGlyHisLysLysLeuLeuGlnLeuLeuThrCysSerSe 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
20732 TTGTCTCCCTCGTCGTCATCATCTCTCGTCAATGTCCTCATCATCGTC 20781
17 TADPAPARGLYHisSerLeuThrAsnSerProLeuAspSerSerC 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
20782 ATCTCTCGTGCCTCGTCGTCGTCCTCGTCGTCGTCGTCGTCGTCGTF 20831

```

[illegible]





## alignment\_block:

US-09-041-994-2\_COPY\_613\_752 x US-08-864-038A-2/rev ..

Align seg 1/1 to reverse of: US-08-864-038A-2 from: 1 to: 3331

```

16 SerSerAspAspArgGlyHisSerSerLeuThrAsnSerProLeuAspSe 32
   ::::::::::: ::::::::::: :::::::::::
520 ACCGCCGAGACCTCCGCCCATCCAGATCCAGATCCAGACCTGCTCTG 471
   ::::::::::: ::::::::::: :::::::::::
32 rSerCysLysGluSerSerValSerValThrSerProSerGlyValSer 49
   ||||| ::::::::::: :::::::::::
470 CTCCTCGCCGCTGCTCAGACACACACACACACCTCCGCGACACACA 421
   ::::::::::: :::::::::::
49 erserThrSerGlyValSerSerThrSerAsnMet..HisGlySerL 65
   ::::::::::: :::::::::::
420 CCAGCACCACCTCCGCGCGCTCCATTCACCGCTCATCCCACTC 371
   :::::::::::
65 euleuGlnGluLysHisArgIleLeuHis..... 74
   :::::::::::
370 CCATTCATCATCGTCATCGTGTGTCATCATGATCATTCATCCAACTC 321
   :::::::::::
75 .....LysLeuLeuGlnAsnGlyAsnSerPr 83
   :::::::::::
320 CACCGCGCGCGCGCTCCGCCGCTCCTCCAGACCGCGAGCTCC 271
   :::::::::::
83 o 83
   |
270 C 270

```

seq\_name: /cgn2\_6/ptodata/1/lna/6A\_COMB.seq:US-08-864-038A-4

## seq\_documentation\_block:

```

; Sequence 4, Application US/08864038A
; Patent No. 6001592
; GENERAL INFORMATION:
; APPLICANT: Kunio NAKASHIMA et al.
; TITLE OF INVENTION: NOVEL POLYPEPTIDE GENE CDNA, VECTOR
; TITLE OF INVENTION: CONTAINING SAID CDNA, HOST CELLS TRANSFORMED WITH SAID
; TITLE OF INVENTION: VECTOR, POLYPEPTIDE PRODUCED THEREBY, METHOD OF PRODUCING
; TITLE OF INVENTION: SAID POLYPEPTIDE, DNA ENCODING SAID POLYPEPTIDE AND ANTIBODY
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 812-5 Hirano
; STREET: Isshinden
; CITY: Tsu-city
; STATE: Mie-prefecture
; COUNTRY: JAPAN
; ZIP: 514-01
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Microsoft Windows 95
; SOFTWARE: Word Perfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/864,038A
; FILING DATE: May 28, 1997
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 8-184459
; FILING DATE: 15-July-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: C. Bruce Hamburg
; REGISTRATION NUMBER: 22,389
; REFERENCE/DOCKET NUMBER: F-5610
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)986-2340
; TELEFAX: (212)953-7733
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3331
; TYPE: nucleic acid
; STRANDEDNESS: single

```

```

; TOPOLOGY: linear
; MOLECULE TYPE: CDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: Pinctada fucata
; CELL TYPE: mantle epithelial cell
; FEATURE:
; NAME/KEY: CDS
; LOCATION: from 50 to 2263
; IDENTIFICATION METHOD: P (by similarity to some other pattern)
; US-08-864-038A-4

```

## alignment\_scores:

Quality:	78.50	Length:	84
Ratio:	1.539	Gaps:	2
Percent Similarity:	60.714	Percent Identity:	29.762

## alignment\_block:

US-09-041-994-2\_COPY\_613\_752 x US-08-864-038A-4/rev ..

Align seg 1/1 to reverse of: US-08-864-038A-4 from: 1 to: 3331

```

16 SerSerAspAspArgGlyHisSerSerLeuThrAsnSerProLeuAspSe 32
   ::::::::::: ::::::::::: :::::::::::
520 ACCGCCGAGACCTCCGCCCATCCAGATCCAGATCCAGACCTGCTCTG 471
   ::::::::::: ::::::::::: :::::::::::
32 rSerCysLysGluSerSerValSerValThrSerProSerGlyValSer 49
   ||||| ::::::::::: :::::::::::
470 CTCCTCGCCGCTGCTCAGACACACACACACACCTCCGCGACACACA 421
   :::::::::::
49 erserThrSerGlyValSerSerThrSerAsnMet..HisGlySerL 65
   ::::::::::: :::::::::::
420 CCAGCACCACCTCCGCGCGCTCCATTCACCGCTCATCCCACTC 371
   :::::::::::
65 euleuGlnGluLysHisArgIleLeuHis..... 74
   :::::::::::
370 CCATTCATCATCGTCATCGTGTGTCATCATGATCATTCATCCAACTC 321
   :::::::::::
75 .....LysLeuLeuGlnAsnGlyAsnSerPr 83
   :::::::::::
320 CACCGCGCGCGCGCTCCGCCGCTCCTCCAGACCGCGAGCTCC 271
   :::::::::::
83 o 83
   |
270 C 270

```

seq\_name: /cgn2\_6/ptodata/1/lna/5A\_COMB.seq:US-08-072-610-1

## seq\_documentation\_block:

```

; Sequence 1, Application US/08072610
; Patent No. 5532133
; GENERAL INFORMATION:
; APPLICANT: Barnwell, John
; TITLE OF INVENTION: Plasmodium vivax Blood Stage Antigens,
; TITLE OF INVENTION: Monoclonal Antibodies, and Diagnostic Assays
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby and Darby
; STREET: 805 Third Ave.
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022-7513
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/072,610
; FILING DATE: 19930602
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

```

NAME: Gogoris, Adda  
REGISTRATION NUMBER: 29,714  
REFERENCE/DOCKET NUMBER: 5986/07686  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)527-7700  
TELEFAX: (212)753-6237  
TELEX: 236687  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3337 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: double  
MOLECULE TYPE: linear  
TOPOLOGY: linear  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Plasmodium vivax  
IMMEDIATE SOURCE:  
CLONE: pVMB3.3.1  
US-08-072-610-1

alignment\_scores:  
Quality: 78.50 Length: 94  
Ratio: 1.377 Gaps: 2  
Percent Similarity: 60.638 Percent Identity: 30.851

## alignment\_block:

US-09-041-994-2\_COPY\_613\_752 x US-08-072-610-1/rev ..

Align seg 1/1 to reverse of: US-08-072-610-1 from: 1 to: 3337

```
14 ThrCysSerSerAspArgGlyHisSerSerLeuThrAsnSerProle 30
||| ||||| ||| ||||| |||||
2368 ACCTCTTCTTCACTACCGCTGCTACTTCTACTACCGCTGCTAC 2319
30 uAspSerSerCysLysGlnSerSerValThrSerProSerGly 47
||||| ||||| ||||| ||||| |||||
2318 CTTCTACTTCT.....TCTACTTCTTCACCTTCTTCACCTTCTCTG 2278
47 aLseSerSerThrSerGlyValSerSerThrSerAsnMetHisGly 63
||||| ||||| ||||| ||||| |||||
2277 GACCTCTTCACCTTCTCTGCTACTTCTTCACCTTCTCTGCTACT 2228
64 SerLeuLeuGlnGlyHisArgIleLeuHisLysLeuLeuGlnAsn 80
||| ||||| ||||| ||||| |||||
2227 TCC.....ACTTCTCTG 2214
80 yAsnSerProAlaGluValAlaLysIleThrAlaGlnAlaThrGly 97
||||| ||||| ||||| ||||| |||||
2213 TACCTCTTCACCTTCTCTGCTACTTCTTCACCTTCTCTGCTACT 2164
97 sPThrSerSerIleThrSerCysGlyAspGly 107
||||| ||||| ||||| ||||| |||||
2163 GTCCTCTTCACCTTCTCTGCTACTTCTGCTGAGT 2132
```

seq\_name: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:US-08-719-822B-1

seq\_documentation\_block:  
Sequence 1, Application US/08719822B  
Patent No. 5874527  
GENERAL INFORMATION:  
APPLICANT: Barnwell, John  
TITLE OF INVENTION: Plasmodium vivax Blood Stage Antigens  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Darby and Darby  
STREET: 805 Third Ave.  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10022-7513

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/719,822B  
FILING DATE: 09/30/96  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Gogoris, Adda  
REGISTRATION NUMBER: 29,714  
REFERENCE/DOCKET NUMBER: 5986/17686US2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)527-7700  
TELEFAX: (212)753-6237  
TELEX: 236687  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3337 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
MOLECULE TYPE: linear  
TOPOLOGY: linear  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Plasmodium vivax  
IMMEDIATE SOURCE:  
CLONE: pVMB3.3.1  
US-08-719-822B-1

alignment\_scores:  
Quality: 78.50 Length: 94  
Ratio: 1.377 Gaps: 2  
Percent Similarity: 60.638 Percent Identity: 30.851

## alignment\_block:

US-09-041-994-2\_COPY\_613\_752 x US-08-719-822B-1/rev ..

Align seg 1/1 to reverse of: US-08-719-822B-1 from: 1 to: 3337

```
14 ThrCysSerSerAspArgGlyHisSerSerLeuThrAsnSerProle 30
||| ||||| ||| ||||| |||||
2368 ACCTCTTCTTCACTACCGCTGCTACTTCTACTTCTACTACCGCTGCTAC 2319
30 uAspSerSerCysLysGlnSerSerValThrSerProSerGly 47
||||| ||||| ||||| ||||| |||||
2318 CTTCTACTTCT.....TCTACTTCTTCACCTTCTTCACCTTCTCTG 2278
47 aLseSerSerThrSerGlyValSerSerThrSerAsnMetHisGly 63
||||| ||||| ||||| ||||| |||||
2277 GACCTCTTCACCTTCTCTGCTACTTCTTCACCTTCTCTGCTACTTCT 2228
64 SerLeuLeuGlnGlyHisArgIleLeuHisLysLeuLeuGlnAsn 80
||| ||||| ||||| ||||| |||||
2227 TCC.....ACTTCTCTG 2214
80 yAsnSerProAlaGluValAlaLysIleThrAlaGlnAlaThrGly 97
||||| ||||| ||||| ||||| |||||
2213 TACCTCTTCACCTTCTCTGCTACTTCTTCACCTTCTCTGCTACTTCTCTG 2164
97 sPThrSerSerIleThrSerCysGlyAspGly 107
||||| ||||| ||||| ||||| |||||
2163 GTCCTCTTCACCTTCTCTGCTACTTCTGCTGAGT 2132
```



This Page Blank (uspto)

```
OM of: US-09-041-994-2_COPY_613_752 to: Pending_Patents_NA_Main:★ out_format : pres
Date: May 1, 2001 8:21 PM
```

**About:** Results were produced by the Gencore software, version 4.5.  
Copyright (c) 1993-2000 compugen Ltd.

```

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool/US09041994/runatc_30042001_165632_19506/app.query.fasta_1.19400
-DB=pending_patents_NA_Main -OFMT=fastlap -SUFFIX=p2n.inpm
-GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPELT=0.000
-LOOPEXT=0.000 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000
-XGAPEXT=0.500 -FGAPOP=6.000 -FGAPEXT=7.000 -XCAPOP=10.000
-YGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1
-MATRIX=blotsum62 -TRANS=human40.cdi -LIST=45 -DOCALLIGN=200
-THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pfs -NORR=ext -MINLEN=0 -MAXLEN=2000000000
-USER=US09041994_@cgn1_1_8354 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXK
-WAIT -THREADS=1

```

```

Query length: 140
Database: Pending_Patents_NA_Main.*
Database sequences: 13168883
Database length: -1691701393
Search time (sec): 4537.790000

```

score\_list:

[illegible]

/cg22_6.p/prodata/1/pma/US096_COMB_seq:us-09-656-992-6922 +	29.50	480.64	1.5e-17
/cg22_6.p/prodata/1/pma/US097A_COMB_seq:us-09-716-972-6417 +	29.50	480.64	1.5e-17
/cg22_6.p/prodata/1/pma/US097B_COMB_seq:us-09-722-559-6496 +	29.50	480.64	1.5e-17
/cg22_6.p/prodata/1/pma/US097C_COMB_seq:us-09-726-178-5097 +	29.50	480.64	1.5e-17
/cg22_6.p/prodata/1/pma/US097B_COMB_seq:us-09-726-178-5097 +	29.50	480.64	1.5e-17
/cg22_6.p/prodata/1/pma/US099_COMB_seq:us-09-041-994-1			

```

seq=documentation_block:
  Sequence 1, Application: US/09041994
  GENERAL INFORMATION:
    APPLICANT: Chen, J. Don
    APPLICANT: Li, Hui
    TITLE OF INVENTION: Transcriptional Coactivator for Nuclear
    TITLE OF INVENTION: Hormone Receptors
    NUMBER OF SEQUENCES: 2
    CORRESPONDENCE ADDRESS:
    ADDRESSEE: Iahive and Cockfield
    STREET: 28 State Street
    City: Boston
    STATE: MA
    COUNTRY: USA
    ZIP: 02109
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/09/041,994
    FILING DATE:
    CLASSIFICATION: 435
  ATTORNEY/AGENT INFORMATION:

```

```

? NAME: liepmann, W. Hugo
? REGISTRATION NUMBER: 20, 407
? REFERENCE/DOCKET NUMBER: UDM-026-1
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 617-227-7400
? TELEFAX: 617-742-4214
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 4496 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 86..4338
? OS-09-041-994-1

```

```
alignment_scores:
  quality: 703.00
  length: 140
  ratio: 5.021
  gaps: 0
  percent similarity: 100.000
  percent identity: 100.000
```

```

alignment_block:
US-09-041-994-2_COPY_613_752 x US-09-041-994-1  ..
Align seg 1/1  to: US-09-041-994-1  from: 1  to: 4496

```

1 LeuGIuSerLysGIyHISLysLysLeuGIuInLeuThrCysSer 17  
1922 TTGGAAAGCAAGCTATATAAAATTACTGAGATTACTACTGTTCTTC 1972			
17 rAspAspArgLyHISerSerLeuThrAsnSerProLeuAspSerSerc 34			
1972 TCATGACCGGGGTCATCTCTCTTGACCAACTCCCCCTAGTTCAAGTT 2022			
34 yLysGIuSerSerValSerValThrSerProSerGIyValSerSerSer 50			
 2022 GTAAGAAATCTTGTGTAGTACCAAGCCCTGTGGAGTCTCCGCCCT 2072

```
51 ThrSerGlyValSerSerThrSerAsnMetHisGlySerLeuLeu1 67
|||||
2072 ACATCTGGAGGAGTATCCTCTACATCCAAATATGCATGGGTCACTGTTTACA 2121
67 nGluysHisArgIleLeuHisLysLeuLeuGlnAsnGlyAsnSerProA 84
|||||
2122 AGAAGAGCCCGGATTTTGCACAAAGTGTCTGCAGAAATGGGAATTCACGAG 2171
84 IagIValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSerSer 100
|||||
2172 CTGAGGTACCCAAAGATTACTGCACAAGCCACTGGGAAAGACACACAGCACT 2221
101 IieThrSerCysGlyAspGlyAsnValValLysGlnGluGlnLeuSerPr 117
|||||
2222 ATAACTCTTGTGGGACGGAATGTTGTCAAGCAGACAGCAGCTAAGTCC 2271
117 oLysLysLysGluAsnAsnAlaLeuLeuArgTyrLeuLeuAspArgAspA 134
|||||
2272 TAAGAAAGAGGAGAAATATGCACTTCTTAGATACCTGCTGCAGACAGGATG 2321
134 sPProSerAspAlaLeuSer 140
|||||
2322 ATCTAGTATGATGCACCTCTCT 2341

seq_name: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:US-09-513-066-1

seq_documentation_block:
: Sequence 1, Application US/09513066
: GENERAL INFORMATION:
: APPLICANT: Chen, J. Don
: APPLICANT: Leo, Christopher
: APPLICANT: Li, Hai
: TITLE OF INVENTION: NOVEL USES FOR THE RAC3 TRANSCRIPTIONAL COACTIVATOR OF
: TITLE OF INVENTION: STEROID NUCLEAR RECEPTORS
: FILE REFERENCE: UMG-026CP
: CURRENT APPLICATION NUMBER: US/09/513,066
: CURRENT FILING DATE: 2000-02-24
: PRIOR APPLICATION NUMBER: USSN 09/041,994
: PRIOR FILING DATE: 1998-03-13
: PRIOR APPLICATION NUMBER: USSN 60/073,674
: PRIOR FILING DATE: 1998-02-04
: NUMBER OF SEQ ID NOS: 37
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 4496
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (86)..(4330)
US-09-513-066-1

alignment_scores:
Quality: 703.00 Length: 140
Ratio: 5.021 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-041-994-2_COPY_613_752 x US-09-513-066-1 ..
Align seg 1/1 to: US-09-513-066-1 from: 1 to: 4496

1 LeuGluSerLysGlyHisLysLysLeuLeuGlnLeuLeuThrCysSerse 17
|||||
1922 TTGGAAGAGCAAGGTCATATAAATTAATGCAATTAATTAATTAATTAATTA 1971
17 rAspAspArgGlyHisSerSerLeuThrAsnSerProLeuAspSerSerSerc 34
|||||
1972 TGATGACCGGGGTCAATTCCTCTTGACCAACTCCCGCTAATTCAGATT 2021
34 yAlysGluSerSerValSerValThrSerProSerGlyValSerSerSer 50
```

```
|||||
2022 GTAAAGAAATCTTCTGTAGTGCACAGCCCTCTGAGACTCTCTCTCT 2071
51 ThrSerGlyValSerSerThrSerAsnMetHisGlySerLeuLeu1 67
|||||
2072 ACATCTGGAGGAGTATCCTCTACATCCAAATATGCATGGGTCACTGTTTACA 2121
67 nGluysHisArgIleLeuHisLysLeuLeuGlnAsnGlyAsnSerProA 84
|||||
2122 AGAAGAGCCCGGATTTTGCACAAAGTGTCTGCAGAAATGGGAATTCACGAG 2171
84 IagIValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSerSer 100
|||||
2172 CTGAGGTACCCAAAGATTACTGCACAAGCCACTGGGAAAGACACACAGCACT 2221
101 IieThrSerCysGlyAspGlyAsnValValLysGlnGluGlnLeuSerPr 117
|||||
2222 ATAACTCTTGTGGGACGGAATGTTGTCAAGCAGACAGCAGCTAAGTCC 2271
117 oLysLysLysGluAsnAsnAlaLeuLeuArgTyrLeuLeuAspArgAspA 134
|||||
2272 TAAGAAAGAGGAGAAATATGCACTTCTTAGATACCTGCTGCAGACAGGATG 2321
134 sPProSerAspAlaLeuSer 140
|||||
2322 ATCTAGTATGATGCACCTCTCT 2341

seq_name: /cgn2_6/ptodata/1/pna/US6006_COMB.seq:US-60-068-511-1

seq_documentation_block:
: Sequence 1, Application US/60068511
: GENERAL INFORMATION:
: APPLICANT: Suen, Chen-Shian
: APPLICANT: Freil, Donald E.
: APPLICANT: Lyttle, Richard C.
: TITLE OF INVENTION: Cloning and Expression of a Nuclear
: TITLE OF INVENTION: Receptor Coactivator Proteins and Uses Thereof
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: American Home Products Corporation
: STREET: One Campus Drive
: CITY: Parsippany
: STATE: New Jersey
: COUNTRY: USA
: ZIP: 07054
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/60/068,511
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Walsh, Andrea C.
: REGISTRATION NUMBER: 34,988
: REFERENCE/DOCKET NUMBER: 97243-00
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 973-683-2169
: TELEFAX: 973-683-4117
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4789 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHEetical: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
: FEATURE:
```

NAME/KEY: CDS  
LOCATION: 185..4750  
US-60-068-511-1

## alignment\_scores:

Quality: 700.00 Length: 140  
Ratio: 5.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 99.286

## alignment\_block:

US-09-041-994-2\_COPY\_613\_752 x US-60-068-511-1

Align seg 1/1 to: US-60-068-511-1 from: 1 to: 4789

```
1 LeuGlSerLysGlyHisLysLysLeuLeuInLeuThrCysSerSe 17
|||||
2021 TTGGAAAGCAAGGTCATATAAAATTACTGCAGTACTTACTGTTCTTC 2070
17 rAspAspArgLysHisSerSerLeuThrAsnSerProLeuAspSerSe 34
|||||
2071 TGATGACCGGGGTCATCTCTTGACCAACTCCCCCTAGATTCAAGTT 2120
34 ySLysGluSerSerValSerValThrSerProSerGlyValSerSerSe 50
|||||
2121 GTAAAGAAATCTTCTGTAGTGTACACACACCCCTCGAGTCTCTCTCT 2170
51 ThrSerGlyGlyValSerSerThrSerAsnMetHisGlySerLeuLeu 67
|||||
2171 ACATCTGGAGAGATCTCTTACATCCAAATATGCATGGTGCATCTGTACA 2220
67 nGluLysHisArgLysLeuHisLysLysLeuLeuGlnAsnGlyAsnSerPro 84
|||||
2221 AGAGAAAGCAGCGGATTTGACAAAGTTGCTGCAGATGGGAATTCACAG 2270
84 lAglValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSerSer 100
|||||
2271 CTGAGGTAGCCAGATTTACTGCAGAGGCACTGGGAAAGACACACGAGT 2320
101 IleThrSerCysGlyAspGlyAsnValValLysGlnGlnLeuSerPr 117
|||||
2321 ATAACTCTTCTGTGGGAGCGGAAATGTTGTCAAGCAGAGCACTAAAGTCC 2370
117 oLysLysLysGlnAsnAsnAlaLeuLeuArgTyrLeuLeuAspArgAsp 134
|||||
2371 TAAGAAAGAGAGAAATATATGCACTTCTTATGATACCTGCTGACAGGAGT 2420
134 sPProSerAspAlaLeuSer 140
|||||
2421 ATCTAGATGACACTCTCT 2440
```

seq\_name: /cgn2\_6/prodata/1/pna/US094\_COMB.seq:US-09-440-612-1

## seq\_documentation\_block:

Sequence 1, Application US/09440612

GENERAL INFORMATION:

APPLICANT: C. Frank Bennett

APPLICANT: Lex M. Cowser

TITLE OF INVENTION: ANTISENSE MODULATION OF SRC-3 EXPRESSION

FILE REFERENCE: RTS-0042

CURRENT APPLICATION NUMBER: US/09/440,612

CURRENT FILING DATE: 1999-11-15

NUMBER OF SEQ ID NOS: 47

SEQ ID NO 1

LENGTH: 6754

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (184)..(4422)

US-09-440-612-1

## alignment\_scores:

Quality: 700.00 Length: 140  
Ratio: 5.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 99.286

## alignment\_block:

US-09-041-994-2\_COPY\_613\_752 x US-09-440-612-1

Align seg 1/1 to: US-09-440-612-1 from: 1 to: 6754

```
1 LeuGlSerLysGlyHisLysLysLeuLeuInLeuThrCysSerSe 17
|||||
2050 TTGGAAAGCAAGGTCATATAAAATTACTGCAGTACTTACTGTTCTTC 2099
17 rAspAspArgLysHisSerSerLeuThrAsnSerProLeuAspSerSe 34
|||||
2100 TGATGACCGGGGTCATCTCTTGACCAACTCCCCCTAGATTCAAGTT 2149
34 ySLysGluSerSerValSerValThrSerProSerGlyValSerSerSe 50
|||||
2150 GTAAAGAAATCTTCTGTAGTGTACACACCCCTCGAGTCTCTCTCTCT 2199
51 ThrSerGlyGlyValSerSerThrSerAsnMetHisGlySerLeuLeu 67
|||||
2200 ACATCTGGAGAGATATCTCTTACATCCAAATATGCATGGTGCATCTGTACA 2249
67 nGluLysHisArgLysLeuHisLysLysLeuLeuGlnAsnGlyAsnSerPro 84
|||||
2250 AGAGAAAGCAGCGGATTTGACAAAGTTGCTGCAGATGGGAATTCACAG 2299
84 lAglValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSerSer 100
|||||
2300 CTGAGGTAGCCAGATTTACTGCAGAGGCACTGGGAAAGACACACGAGT 2349
101 IleThrSerCysGlyAspGlyAsnValValLysGlnGlnLeuSerPr 117
|||||
2350 ATAACTTCTTGTGGGAGCGGAAATGTTGTCAAGCAGAGCACTAAAGTCC 2399
117 oLysLysLysGlnAsnAsnAlaLeuLeuArgTyrLeuLeuAspArgAsp 134
|||||
2400 TAAGAAAGAGAGAAATATATGCACTTCTTATGATACCTGCTGACAGGAGT 2449
134 sPProSerAspAlaLeuSer 140
|||||
2450 ATCTAGATGACACTCTCT 2469
```

seq\_name: /cgn2\_6/prodata/1/pna/US094\_COMB.seq:US-09-440-612-3

## seq\_documentation\_block:

Sequence 3, Application US/09440612A

GENERAL INFORMATION:

APPLICANT: C. Frank Bennett

APPLICANT: Lex M. Cowser

TITLE OF INVENTION: ANTISENSE MODULATION OF SRC-3 EXPRESSION

FILE REFERENCE: RTS-0042

CURRENT APPLICATION NUMBER: US/09/440,612A

CURRENT FILING DATE: 1999-11-15

NUMBER OF SEQ ID NOS: 49

SEQ ID NO 3

LENGTH: 6754

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (184)..(4422)

US-09-440-612-3

## alignment\_scores:

Quality: 700.00 Length: 140  
Ratio: 5.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 99.286

alignment\_block:  
US-09-041-994-2\_COPY\_613\_752 x US-09-440-612-3 ..

Align seg 1/1 to: US-09-440-612-3 from: 1 to: 6754

```
1 LeuGlSerLysGlyHisLysLysLeuGlnLeuLeuThrCysSerse 17
|||||
2050 TTGGAAAGCAAGGTCATAAAAAATTACTGCACTTACTTACTGTTCTTC 2099
17 RASPAAPATGGLYHISERSELEUTHRASNSERPROLEUASPSESR 34
|||||
2100 TGATGACCGGGGTCATTCTCTTGACCAACTCCCCCTGATTCAGATT 2149
34 YLSLSEUSERSERVALSERVALTHRSERPROSERGLYVALSERSE 50
|||||
2150 GTAAGAAATCTTCTGTAGTGTACACAGCCCCCTGTGAGTCTCTCTCT 2199
51 THRSERGLYVALSERSETRHRASNMETHISGLYSERLEUENGL 67
|||||
2200 ACATCTGGAGAGATATCTTACATCCAAATATGATGGGTCACTGTACA 2249
67 NGIULYSHISARGILEUHIISLysLeuGlnAsnGlyAsnSerPro 84
|||||
2250 AGAAGAACACCGGATTTTGCACAAGTTGTCGACGAATGGGAATTCAC 2299
84 IAGLUVALALALYSILETHRALAGLINALATHRGLYLSASPTHRSE 100
|||||
2300 CTGAGGTAGCCAGATTACTGCGAAGCCACTGGAAAGACACACAGT 2349
101 ILETHRSECYSGLYASPGLYASNVALYALYSLGlnGlnLeuSerP 117
|||||
2350 ATAACTCTTGTGTGGGACGGAATGTTGTCAACAGACAGACGTAAGTC 2399
117 OLYSLYSLYGLUASNASNALALEUENARGTYRLEULEUASPARG 134
|||||
2400 TAAGAAGAAAGGAATATATGCACTTCTTAGATACCTGCTGACAGGATG 2449
134 SPROSERASPALALEUSER 140
|||||
2450 ATCTAGTAGTCACCTCTCT 2469
```

seq\_name: /cgn2\_6/ptodata/1/pna/US091\_COMB.seq:US-09-125-635-1

seq\_documentation\_block:  
; Sequence 1, Application US/09125635  
; GENERAL INFORMATION:  
; APPLICANT: THE UNITED STATES OF AMERICA represented by THE SE  
; TITLE OF INVENTION: AIB1, A novel steroid receptor co-activator  
; FILE REFERENCE: 49944  
; CURRENT APPLICATION NUMBER: US/09/125,635  
; CURRENT FILING DATE: 1998-08-21  
; PRIOR APPLICATION NUMBER: 60/049,728  
; PRIOR FILING DATE: 1997-06-17  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 6835  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (201)..(4463)  
US-09-125-635-1

alignment\_scores:  
Quality: 700.00 Length: 140  
Ratio: 5.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 99.286

alignment\_block:  
US-09-041-994-2\_COPY\_613\_752 x US-09-125-635-1 ..

Align seg 1/1 to: US-09-125-635-1 from: 1 to: 6835

```
1 LeuGlSerLysGlyHisLysLysLeuGlnLeuLeuThrCysSerse 17
|||||
2037 TTGGAAAGCAAGGTCATAAAAAATTACTGCACTTACTTACTGTTCTTC 2086
17 RASPAAPATGGLYHISERSELEUTHRASNSERPROLEUASPSESR 34
|||||
2087 TGATGACCGGGGTCATTCTCTTGACCAACTCCCCCTGATTCAGATT 2136
34 YLSLSEUSERSERVALSERVALTHRSERPROSERGLYVALSERSE 50
|||||
2137 GTAAGAAATCTTCTGTAGTGTACACAGCCCCCTGTGAGTCTCTCTCT 2186
51 THRSERGLYVALSERSETRHRASNMETHISGLYSERLEUENGL 67
|||||
2187 ACATCTGGAGAGATATCTTACATCCAAATATGATGGGTCACTGTACA 2236
67 NGIULYSHISARGILEUHIISLysLeuGlnAsnGlyAsnSerPro 84
|||||
2237 AGAAGAACACCGGATTTTGCACAAGTTGTCGACGAATGGGAATTCAC 2286
84 IAGLUVALALALYSILETHRALAGLINALATHRGLYLSASPTHRSE 100
|||||
2287 CTGAGGTAGCCAGATTACTGCGAAGCCACTGGAAAGACACACAGT 2336
101 ILETHRSECYSGLYASPGLYASNVALYALYSLGlnGlnLeuSerP 117
|||||
2337 ATAACTCTTGTGTGGGACGGAATGTTGTCAACAGACAGACGTAAGTC 2386
117 OLYSLYSLYGLUASNASNALALEUENARGTYRLEULEUASPARG 134
|||||
2387 TAAGAAGAAAGGAATATATGCACTTCTTAGATACCTGCTGACAGGATG 2436
134 SPROSERASPALALEUSER 140
|||||
2437 ATCTAGTAGTCACCTCTCT 2456
```

seq\_name: /cgn2\_6/ptodata/1/pna/US093\_COMB.seq:US-09-359-922-3962

seq\_documentation\_block:  
; Sequence 3962, Application US/09359922  
; GENERAL INFORMATION:  
; APPLICANT: Hysq, Inc.  
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA  
; TITLE OF INVENTION: LIBRARIES  
; FILE REFERENCE: 20411-752CON1  
; CURRENT APPLICATION NUMBER: US/09/359,922  
; CURRENT FILING DATE: 1999-07-22  
; EARLIER APPLICATION NUMBER: US 09/205,155  
; EARLIER FILING DATE: 1998-12-03  
; NUMBER OF SEQ ID NOS: 13203  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3962  
; LENGTH: 6855  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)..(6855)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-359-922-3962

alignment\_scores:  
Quality: 700.00 Length: 140  
Ratio: 5.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 99.286

alignment\_block:  
US-09-041-994-2\_COPY\_613\_752 x US-09-359-922-3962 ..

Align seg 1/1 to: US-09-359-922-3962 from: 1 to: 6855



```

1 LeuGIserLysGlyHisLysLysLeuLeuGlnLeuLeuThrCysSer 17
|||||
2037 TTGAAAGCAAGAGTCATATAAAATTACTGCAGTACTTACCTGTTCTTC 2086

17 RASPAAPRGlyHisSerSerLeuThrAsnSerProLeuAspSerSerc 34
|||||
2087 TGATGACCGGGGTCATTCCTTGCACCAACTCCCTTACATTCAGTT 2136

34 YSLysGlySerSerValSerValThrSerProSerGlyValSerSer 50
|||||
2137 GTAAGAAATCTTCTGTAGTGTACACAGCCCTCTGGAGTCTCTCTCT 2186

51 ThrSerGlyValSerSerThrSerAsnMetHisGlySerLeuLeuG 67
|||||
2187 ACATCTGGAGAGATACCTCTACATCCAAATGATGATGGCTACGTATCA 2236

67 nGluLysHisArgIleLeuHisLysLeuLeuGlnAsnGlyAsnSerPro 84
|||||
2237 AGAAGACACCGCATTTTGCACAAAGTTCTGCAGAAATGGAAATTCAC 2286

84 IagIuValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSer 100
|||||
2287 CTGAGGTAGCCCAAGATTACTGCAGAGCCACTGGGAAAGACACAGCAG 2336

101 IleThrSerCysGlyAspGlyAsnValValLysGlnGlnLeuSerPr 117
|||||
2337 ATAACCTTCTTGTGGGAGCAAAATGTTCTCAAGACGAGCAGCTAAAG 2386

117 OlyLysLysGlyLysAsnAlaLeuLeuArgTyrLeuLeuAspArgAs 134
|||||
2387 TAAAGAGAGAGAGATATATGCACTTCTTAGATACTGCTGCAGACGAG 2436

134 sPProSerAspAlaLeuSer 140
|||||
2437 ATCCTAGTATGACACTCTCT 2456

seq_name: /cgn2_6/ptodata/1/pna/US093_COMB.seq:us-09-359-922-3962

seq_documentation_block:
: Sequence 3962, Application US/09359922A
: GENERAL INFORMATION:
: APPLICANT: Ieshkowitz, Dena
: APPLICANT: Liu, Jin
: TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
: FILE REFERENCE: 20411-752CON1
: CURRENT APPLICATION NUMBER: US/09/359,922A
: CURRENT FILING DATE: 1999-07-22
: EARLIER APPLICATION NUMBER: US 09/205,155
: EARLIER FILING DATE: 1998-12-03
: EARLIER APPLICATION NUMBER: US 09/034,341
: EARLIER FILING DATE: 1998-02-13
: NUMBER OF SEQ ID NOS: 13203
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 3962
: LENGTH: 6855
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)-(6855)
: OTHER INFORMATION: n = A,T,C or G
US-09-359-922-3962

```

```

alignment_scores:
  Quality: 700.00      Length: 140
  Ratio: 5.000        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 99.286

alignment_block:
US-09-041-994-2_copy_613_752 x US-09-359-922-3962

```

```

Align seg 1/1 to: US-09-359-922-3962 from: 1 to: 6855

1 LeuGIserLysGlyHisLysLysLeuLeuGlnLeuLeuThrCysSer 17
|||||
2037 TTGAAAGCAAGAGTCATATAAAATTACTGCAGTACTTACCTGTTCTTC 2086

17 RASPAAPRGlyHisSerSerLeuThrAsnSerProLeuAspSerSerc 34
|||||
2087 TGATGACCGGGGTCATTCCTTGCACCAACTCCCTTACATTCAGTT 2136

34 YSLysGlySerSerValSerValThrSerProSerGlyValSerSer 50
|||||
2137 GTAAGAAATCTTCTGTAGTGTACACAGCCCTCTGGAGTCTCTCTCT 2186

51 ThrSerGlyValSerSerThrSerAsnMetHisGlySerLeuLeuG 67
|||||
2187 ACATCTGGAGAGATACCTCTACATCCAAATGATGATGGCTACGTATCA 2236

67 nGluLysHisArgIleLeuHisLysLeuLeuGlnAsnGlyAsnSerPro 84
|||||
2237 AGAAGACACCGCATTTTGCACAAAGTTCTGCAGAAATGGAAATTCAC 2286

84 IagIuValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSer 100
|||||
2287 CTGAGGTAGCCCAAGATTACTGCAGAGCCACTGGGAAAGACACAGCAG 2336

101 IleThrSerCysGlyAspGlyAsnValValLysGlnGlnLeuSerPr 117
|||||
2337 ATAACCTTCTTGTGGGAGCAAAATGTTCTCAAGACGAGCAGCTAAAG 2386

117 OlyLysLysGlyLysAsnAlaLeuLeuArgTyrLeuLeuAspArgAs 134
|||||
2387 TAAAGAGAGAGAGATATATGCACTTCTTAGATACTGCTGCAGACGAG 2436

134 sPProSerAspAlaLeuSer 140
|||||
2437 ATCCTAGTATGACACTCTCT 2456

seq_name: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:US-60-209-009-198

seq_documentation_block:
: Sequence 198, Application US/60209009
: GENERAL INFORMATION:
: APPLICANT: Faris, Mary
: APPLICANT: Pearson, Cecelia I.
: TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
: FILE REFERENCE: PA-0027 P
: CURRENT APPLICATION NUMBER: US/60/209,009
: CURRENT FILING DATE: 2000-06-01
: NUMBER OF SEQ ID NOS: 501
: SOFTWARE: PERL Program
: SEQ ID NO 198
: LENGTH: 7116
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No: 1094199.1
: NAME/KEY: unsure
: LOCATION: 3941-3993, 5899-5939, 6951
: OTHER INFORMATION: a, t, c, g, or other
US-60-209-009-198

```

```

alignment_scores:
  Quality: 700.00      Length: 140
  Ratio: 5.000        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 99.286

alignment_block:
US-09-041-994-2_copy_613_752 x US-60-209-009-198

```

Align seg 1/1 to: US-60-209-009-198 from: 1 to: 7116

```
1 LeuGluSerLysGlyHisLysLysLeuGlnLeuLeuThrcysSerse 17
|||||
2037 TTGGAAAGCAAAAGGTCATAAAAAATTACTGACAGTACTTACCTGTTCTTCT 2086
17 rAspaSPaRgGlyHisSerSerLeuThrAsnSerProLeuAspSerSerc 34
|||||
2087 TGATGACCGGGGTCATTCTCTTGACCAATCCGCCCTAGATTCAAGTT 2136
34 yLysGluSerSerValSerValThrSerProSerGlyValSerSerSer 50
|||||
2137 GTAAGAAATCTTCTGTAGTGTACACAGCCCTCTGAGTCTCTCTCT 2186
51 ThrSerGlyGlyValSerSerThrSerAsnMetHisGlySerLeuGlu 67
|||||
2187 ACATCTGGAGGAGTATCCTCTACATCCAAATATGCAATGGCTCATGTTACA 2236
67 nGluLysHisArgGlyLeuHisLysLeuGlnAsnGlyAsnSerProA 84
|||||
2237 AGAGAAAGCACCGGATTTGACAAAGTTGCTGCAGAAATGGAAATTCACAG 2286
84 lAgLysValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSerSer 100
|||||
2287 CTGAGGTACCAAGATTACTGCAGAGCCACTGGAAAGACACACACAGT 2336
101 IleThrSerCysGlyAspGlyAsnValValLysGlnGluGlnLeuSerP 117
|||||
2337 ATAACTTCTTGTGGGACGGAATGTTGTCAAGCAGACAGCAGCTAAGTCC 2386
117 oLysLysLysGluAsnAsnAlaLeuLeuArgTyrLeuLeuAspArgAspA 134
|||||
2387 TAAGAAAGAGAGAAATATGCACTTCTTAGATACCTGCTGACAGGAGGATG 2436
134 sPProSerAspAlaLeuSer 140
|||||
2437 ATCCTAGTGTATGCACCTCT 2456
```

seq\_name: /cgn2\_6/ptodata/1/pna/US6021\_COMB.seq:US-60-213-360-7991

```
seq_documentation_block:
: Sequence 7991, Application US/60213360
: GENERAL INFORMATION:
: APPLICANT: Morris, MacDonald
: APPLICANT: Lai, Preclt
: TITLE OF INVENTION: Method for the identification of sequence polymorphisms using
: TITLE OF INVENTION: Polynucleotide sequence databases, and single nucleotide polymor
: FILE REFERENCE: GX-0014 P
: CURRENT APPLICATION NUMBER: US/60/213, 360
: CURRENT FILING DATE: 2000-06-21
: NUMBER OF SEQ ID NOS: 8347
: SOFTWARE: PERL Program
: SEQ ID NO 7991
: LENGTH: 7116
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc.feature
: OTHER INFORMATION: Incyte ID No: 1094199.1
: NAME/KEY: unsure
: LOCATION: 3941-3993, 5899-5939, 6951
: OTHER INFORMATION: a, t, c, g, or other
US-60-213-360-7991
```

alignment\_scores:

Quality: 700.00 Length: 140  
Ratio: 5.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 99.286

alignment\_block:

US-09-041-994-2\_COPY\_613\_752 x US-60-213-360-7991 ..

Align seg 1/1 to: US-60-213-360-7991 from: 1 to: 7116

```
1 LeuGluSerLysGlyHisLysLysLeuGlnLeuLeuThrcysSerse 17
|||||
2037 TTGGAAAGCAAAAGGTCATAAAAAATTACTGACAGTACTTACCTGTTCTTCT 2086
17 rAspaSPaRgGlyHisSerSerLeuThrAsnSerProLeuAspSerSerc 34
|||||
2087 TGATGACCGGGGTCATTCTCTTGACCAATCCGCCCTAGATTCAAGTT 2136
34 yLysGluSerSerValSerValThrSerProSerGlyValSerSerSer 50
|||||
2137 GTAAGAAATCTTCTGTAGTGTACACAGCCCTCTGAGTCTCTCTCTCT 2186
51 ThrSerGlyGlyValSerSerThrSerAsnMetHisGlySerLeuGlu 67
|||||
2187 ACATCTGGAGGAGTATCCTCTACATCCAAATATGCAATGGCTCATGTTACA 2236
67 nGluLysHisArgGlyLeuHisLysLeuGlnAsnGlyAsnSerProA 84
|||||
2237 AGAGAAAGCACCGGATTTGACAAAGTTGCTGCAGAAATGGAAATTCACAG 2286
84 lAgLysValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSerSer 100
|||||
2287 CTGAGGTACCAAGATTACTGCAGAGCCACTGGAAAGACACACACAGT 2336
101 IleThrSerCysGlyAspGlyAsnValValLysGlnGluGlnLeuSerP 117
|||||
2337 ATAACTTCTTGTGGGACGGAATGTTGTCAAGCAGACAGCAGCTAAGTCC 2386
117 oLysLysLysGluAsnAsnAlaLeuLeuArgTyrLeuLeuAspArgAspA 134
|||||
2387 TAAGAAAGAGAGAAATATGCACTTCTTAGATACCTGCTGACAGGAGGATG 2436
134 sPProSerAspAlaLeuSer 140
|||||
2437 ATCCTAGTGTATGCACCTCT 2456
```

seq\_name: /cgn2\_6/ptodata/1/pna/US096C\_COMB.seq:US-09-652-123-8947

```
seq_documentation_block:
: Sequence 8947, Application US/09652123
: GENERAL INFORMATION:
: APPLICANT: Holtzman, Douglas A.
: TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
: TITLE OF INVENTION: THEREFOR
: FILE REFERENCE: 1600,1186-001
: CURRENT APPLICATION NUMBER: US/09/652,123
: CURRENT FILING DATE: 2000-08-30
: PRIOR APPLICATION NUMBER: 60/151,135
: PRIOR FILING DATE: 1999-08-30
: NUMBER OF SEQ ID NOS: 9796
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 8947
: LENGTH: 7888
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-652-123-8947
```

alignment\_scores:

Quality: 700.00 Length: 140  
Ratio: 5.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 99.286

alignment\_block:

US-09-041-994-2\_COPY\_613\_752 x US-09-652-123-8947 ..

Align seg 1/1 to: US-09-652-123-8947 from: 1 to: 7888

1 LeuGluSerLysGlyHisLysLysLeuGlnLeuLeuThrcysSerse 17

```

2048 TTGAAAGCAAGCTCATPAAAAAATTACTGCAGTACTACTCTTCTC 2097
17 rAspAspArgGlyHisSerSerLeuThrAsnSerProLeuAspSerSerC 34
2148 GTTAAAGAAATCTTCTGTAGTGTACACAGCCCTCTGGAGTCTCTCTC 2197
2098 TGATGACCGGGGTCATCTCTTCTTGCACCACTCCCTTACATTCAGATT 2147
34 yslYsgIuSerSerValSerValThrSerProSerGlyValSerSerSer 50
2148 GTTAAAGAAATCTTCTGTAGTGTACACAGCCCTCTGGAGTCTCTCTC 2197
51 ThrSerGlyValSerSerThrSerAsnMetHisGlySerLeuLeuG 67
2198 ACATCTGGAGAGATATCTCTACATCCAAATTCATGGTCTACCTTACA 2247
67 nGlulysHisArgIleLeuHisLysLeuLeuGlnAsnGlyAsnSerProA 84
2248 AGAAGACACCGGATTTTGCACAAATGTCGCAAGATGGGAATTCACAC 2297
84 lAgIuValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSerSer 100
2298 CTGAGTAGCCAGATTAATCTGACAGACCCACTGGGAAAGACACACAG 2347
101 IleThrSerCysGlyAspGlyAsnValValLysGlnGluGlnLeuSerPr 117
2348 ATAACTTCTTGTGGGAGCGGAATGTTGTCAAGCAGAGCAGCTAAGTCC 2397
117 OlyLysLysGlnAsnAsnAlaLeuLeuArgTyrLeuLeuAspArgAspA 134
2398 TAAAGAAAGAGAGATATATGCACTTCTTAAATACCTCTGACAGGATG 2447
134 sPProSerAspAlaLeuSer 140
2448 ATCTAGTATGACACTCTCT 2467

```

seq\_name: /cgn2\_6/plodata/1/pna/US096C\_COMB.seq:US-09-652-127-9831

```

seq_documentation_block:
; Sequence 9831, Application US/09652127
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1183-001
; CURRENT APPLICATION NUMBER: US/09/652,127
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,134
; NUMBER OF SEQ ID NOS: 10475
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 9831
; LENGTH: 7888
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-127-9831

```

alignment\_scores: Quality: 700.00 Length: 140  
Ratio: 5.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 99.286

alignment\_block: US-09-041-994-2\_COPY\_613\_752 x US-09-652-127-9831

Align seg 1/1 to: US-09-652-127-9831 from: 1 to: 7888

```

1 LeuGluSerLysGlyHisLysLysLeuLeuGlnLeuLeuThrCysSerS 17
2048 TTGAAAGCAAGCTCATPAAAAAATTACTGCAGTACTACTCTTCTC 2097
17 rAspAspArgGlyHisSerSerLeuThrAsnSerProLeuAspSerSerC 34

```

```

2098 TGATGACCGGGGTCATCTCTTCTTGCACCACTCCCTTACATTCAGTT 2147
34 yslYsgIuSerSerValSerValThrSerProSerGlyValSerSerSer 50
2148 GTTAAAGAAATCTTCTGTAGTGTACACAGCCCTCTGGAGTCTCTCTC 2197
51 ThrSerGlyValSerSerThrSerAsnMetHisGlySerLeuLeuG 67
2198 ACATCTGGAGAGATATCTCTACATCCAAATTCATGGTCTACCTTACA 2247
67 nGlulysHisArgIleLeuHisLysLeuLeuGlnAsnGlyAsnSerProA 84
2248 AGAAGACACCGGATTTTGCACAAATGTCGCAAGATGGGAATTCACAC 2297
84 lAgIuValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSerSer 100
2298 CTGAGTAGCCAGATTAATCTGACAGACCCACTGGGAAAGACACACAG 2347
101 IleThrSerCysGlyAspGlyAsnValValLysGlnGluGlnLeuSerPr 117
2348 ATAACTTCTTGTGGGAGCGGAATGTTGTCAAGCAGAGCAGCTAAGTCC 2397
117 OlyLysLysGlnAsnAsnAlaLeuLeuArgTyrLeuLeuAspArgAspA 134
2398 TAAAGAAAGAGAGATATATGCACTTCTTAAATACCTCTGACAGGATG 2447
134 sPProSerAspAlaLeuSer 140
2448 ATCTAGTATGACACTCTCT 2467

```

seq\_name: /cgn2\_6/plodata/1/pna/US096C\_COMB.seq:US-09-652-355-11058

```

seq_documentation_block:
; Sequence 11058, Application US/09652355
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1192-001
; CURRENT APPLICATION NUMBER: US/09/652,355
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,136
; NUMBER OF SEQ ID NOS: 11227
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 11058
; LENGTH: 7888
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-355-11058

```

alignment\_scores: Quality: 700.00 Length: 140  
Ratio: 5.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 99.286

alignment\_block: US-09-041-994-2\_COPY\_613\_752 x US-09-652-355-11058

Align seg 1/1 to: US-09-652-355-11058 from: 1 to: 7888

```

1 LeuGluSerLysGlyHisLysLysLeuLeuGlnLeuLeuThrCysSerS 17
2048 TTGAAAGCAAGCTCATPAAAAAATTACTGCAGTACTACTCTTCTC 2097
17 rAspAspArgGlyHisSerSerLeuThrAsnSerProLeuAspSerSerC 34
2098 TGATGACCGGGGTCATCTCTTCTTGCACCACTCCCTTACATTCAGATT 2147
34 yslYsgIuSerSerValSerValThrSerProSerGlyValSerSerSer 50
2148 GTTAAAGAAATCTTCTGTAGTGTACACAGCCCTCTGGAGTCTCTCTC 2197

```

```
51 ThrSerGlyGlyValSerSerThrSerAsnMetHisGlySerLeuLeu1 67
|||||
2198 ACATCTGGAGGAGTATCTCTACATCCAAATATGCAATGGGTCACTGTACA 2247
67 nCluYsNHsArGlyLeuNHsLysLeuLeuGlnAsnGlyAsnSerProA 84
|||||
2248 AGAGAAACCCGATTTTGCACAAAGTTCCTGCAGAAATGGGAATTCACCAAG 2297
84 laGlValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSerSer 100
|||||
2298 CTGAGGTACCAAGATTTACTGCAGAAAGCCACTGGGAAAGACACAGCACT 2347
101 IleThrSerCysGlyAspGlyAsnValValLysGlnGlnLeuSerPr 117
|||||
2348 ATAACTTCTTGTGGGACGGAATGTTGTCAAGCAGAGCACGACTAAGTCC 2397
117 oLysLysLysGluAsnAsnAlaLeuLeuArgTyrLeuLeuAspArgAspA 134
|||||
2398 TAAGAAAGAGGAGAAATATGCACTTCTTAGATACCTGCTGCAGAGGAGT 2447
134 sPProSerAspAlaLeuSer 140
|||||
2448 ATCTAGTATGACTCTCT 2467
```

seq\_name: /cgn2\_6/ptodata/1/pna/US096D\_COMB.seq:US-09-663-693-1028

```
seq_documentation_block:
: Sequence 1028, Application US/09663693
: GENERAL INFORMATION:
: APPLICANT: Gearing, David P.
: TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
: FILE OF INVENTION: THEREFOR
: FILE REFERENCE: 1600.1195-001
: CURRENT APPLICATION NUMBER: US/09/663,693
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: 60/154,986
: PRIOR FILING DATE: 1999-09-21
: NUMBER OF SEQ ID NOS: 1340
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1028
: LENGTH: 7888
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-663-693-1028
```

alignment\_scores:  
Quality: 700.00 Length: 140  
Ratio: 5.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 99.286

alignment\_block:  
US-09-041-994-2\_COPY\_613\_752 x US-09-663-693-1028 ..

Align seg 1/1 to: US-09-663-693-1028 from: 1 to: 7888

```
1 LeuGluSerLysGlyNHsLysLysLeuLeuGlnLeuLeuThrCysSerSe 17
|||||
2048 TTGGAAAGCAAGGTCATATAAAATTAATGAGTAACTTAACCTGTTCTTC 2097
17 rAspAspArgGlyNHsSerSerLeuThrAsnSerProLeuAspSerSerC 34
|||||
2098 TGATACCGGGGTCAATCTCTCTTGACCAACTCCCCCTAGATTCAAGTT 2147
34 yAlysGluSerSerValSerValThrSerProSerGlyValSerSerSer 50
|||||
2148 GTAAAGAAATCTTCTGTAGTGCACACAGCCCTCGAGATCTCTCTCT 2197
51 ThrSerGlyGlyValSerSerThrSerAsnMetHisGlySerLeuLeu1 67
|||||
2198 ACATCTGGAGGAGTATCTCTACATCCAAATATGCAATGGGTCACTGTACA 2247
```

```
67 nCluYsNHsArGlyLeuNHsLysLeuLeuGlnAsnGlyAsnSerProA 84
|||||
2248 AGAGAAACCCGATTTTGCACAAAGTTCCTGCAGAAATGGGAATTCACCAAG 2297
84 laGlValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSerSer 100
|||||
2298 CTGAGGTACCAAGATTTACTGCAGAAAGCCACTGGGAAAGACACAGCACT 2347
101 IleThrSerCysGlyAspGlyAsnValValLysGlnGlnLeuSerPr 117
|||||
2348 ATAACTTCTTGTGGGACGGAATGTTGTCAAGCAGAGCACGACTAAGTCC 2397
117 oLysLysLysGluAsnAsnAlaLeuLeuArgTyrLeuLeuAspArgAspA 134
|||||
2398 TAAGAAAGAGGAGAAATATGCACTTCTTAGATACCTGCTGCAGAGGAGT 2447
134 sPProSerAspAlaLeuSer 140
|||||
2448 ATCTAGTATGACTCTCT 2467
```

seq\_name: /cgn2\_6/ptodata/1/pna/US096E\_COMB.seq:US-09-698-014-5135

```
seq_documentation_block:
: Sequence 5135, Application US/09698014
: GENERAL INFORMATION:
: APPLICANT: Gearing, David P.
: APPLICANT: Holtzman, Douglas A.
: TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
: FILE OF INVENTION: THEREFOR
: FILE REFERENCE: 1600.2014-001
: CURRENT APPLICATION NUMBER: US/09/698,014
: PRIOR FILING DATE: 2000-10-27
: PRIOR APPLICATION NUMBER: 60/162,363
: PRIOR FILING DATE: 1999-10-29
: NUMBER OF SEQ ID NOS: 6098
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 5135
: LENGTH: 7888
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-698-014-5135
```

alignment\_scores:  
Quality: 700.00 Length: 140  
Ratio: 5.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 99.286

alignment\_block:  
US-09-041-994-2\_COPY\_613\_752 x US-09-698-014-5135 ..

Align seg 1/1 to: US-09-698-014-5135 from: 1 to: 7888

```
1 LeuGluSerLysGlyNHsLysLysLeuLeuGlnLeuLeuThrCysSerSe 17
|||||
2048 TTGGAAAGCAAGGTCATATAAAATTAATGAGTAACTTAACCTGTTCTTC 2097
17 rAspAspArgGlyNHsSerSerLeuThrAsnSerProLeuAspSerSerC 34
|||||
2098 TGATACCGGGGTCAATCTCTCTTGACCAACTCCCCCTAGATTCAAGTT 2147
34 yAlysGluSerSerValSerValThrSerProSerGlyValSerSerSer 50
|||||
2148 GTAAAGAAATCTTCTGTAGTGCACACAGCCCTCGAGATCTCTCTCT 2197
51 ThrSerGlyGlyValSerSerThrSerAsnMetHisGlySerLeuLeu1 67
|||||
2198 ACATCTGGAGGAGTATCTCTACATCCAAATATGCAATGGGTCACTGTACA 2247
67 nCluYsNHsArGlyLeuNHsLysLeuLeuGlnAsnGlyAsnSerProA 84
|||||
2248 AGAGAAACCCGATTTTGCACAAAGTTCCTGCAGAAATGGGAATTCACCAAG 2297
```

```
84  IagluValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSerSer 100
|||||
2298 CTGAGGTAGCCAGATTACTGCAGAACCCACTGGGAAGACACACGACGT 2347
|||||
101  IleThrSerCysGlyAspGlyAsnValValLysGlnGlnLeuSerPr 117
|||||
2348 ATAACTTCTGTGGGACGGAAATGTTGTCAGCAGCAGCAGCTAAGTCC 2397
|||||
117  OlyLysLysGluAsnAsnAlaLeuLeuArgTyrLeuLeuAspArgAspA 134
|||||
2398 TAAGAAAGAGAGAGATATATGCACTTCTTAGATACCTGCGACACGGATG 2447
|||||
134  sPProSerAspAlaLeuSer 140
|||||
2448 ATCCTAGTGATGCACCTCTCT 2467
```

This Page Blank (uspto)

OM of: US-09-041-994-2\_COPY\_613\_752 to: Pending\_Patents\_NA\_New:\* out\_format : pfs  
Date: May 1, 2001 9:18 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

## Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-O=/cgn2\_1/USPRO.spool/US09041994/tunat\_30042001.165632.19523/app-query.fasta.1.1940  
-DB=Pending\_Patents\_NA\_New -OEM=fastap -SUFFIX=p2n.rnpn  
-GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000  
-LOOEXT=0.000 -OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000  
-XGAEXT=0.500 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000  
-YGAEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1  
-MATRIX=blomsun62 -TRANS=human40.cdi -LIST=45 -DOCALLGN=200  
-THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFM=pfs -NORM=ext -MINLEN=0 -MAXLEN=200000000  
-USER=US09041994\_@cgn1\_1.175 -NCPU=6 -ICPU=3 -LONGLOG -NO\_XLPRX  
-WAIT -THREADS=1

## Search information block:

Query: US-09-041-994-2\_COPY\_613\_752  
Query length: 140  
Database: Pending\_Patents\_NA\_New:\*  
Database sequences: 201016  
Database length: 146721182  
Search time (sec): 114.180000

## Score list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-196-296B-1 +		700.00	1304.56	3.0e-65	47
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12207 -		99.00	102.17	280.06	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12210 -		87.50	83.52	3.1e+03	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12210 -		86.00	80.66	4.4e+03	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12210 -		85.50	131.12	6.83	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12214 -		85.50	84.48	2.7e+03	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12212 -		85.00	87.55	1.8e+03	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12218 +		83.00	74.93	9.1e+03	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12212 +		82.00	81.82	3.8e+03	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12208 +		80.00	75.42	8.6e+03	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-956-171C-480 +		78.50	116.26	45.93	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-956-171C-420 +		78.50	109.22	113.41	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-739-449-218 +		78.00	72.84	1.2e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12216 -		78.00	67.00	2.5e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12210 -		77.00	66.76	2.5e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-956-171C-296B +		77.00	133.33	5.14	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-813-206-736 +		77.00	132.22	5.93	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-281-593-108 +		76.50	117.56	38.88	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12213 -		76.50	66.22	2.7e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		76.00	61.69	4.7e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-540-212A-32765 +		75.50	141.43	1.82	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-801-833-6569 +		75.50	124.07	16.88	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-811-380-681 +		75.00	116.43	44.95	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12211 +		74.50	55.43	3.1e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12207 +		74.50	55.43	9.8e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-248-505-604 +		74.00	88.58	1.6e+03	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-801-833-8092 -		73.50	117.13	41.08	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-423-844-169 +		73.00	114.83	55.22	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-423-844-169 +		73.00	112.66	20.23	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-956-171C-4457 +		73.00	132.49	5.73	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.50	60.01	5.8e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.50	56.92	8.3e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12216 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12206 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		73.00	68.83	2.0e+04	
/cgn2_6/p/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-122					

```

seq_name: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-032-12207
seq_documentation_block:
; Sequence 12207, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; TITLE OF INVENTION: Transcriptome
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12207
; LENGTH: 1531974
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-12207

alignment_scores:
Quality: 99.00 Length: 120
Ratio: 1.414 Gaps: 4
Percent Similarity: 58.333 Percent Identity: 26.667

alignment_block:
US-09-041-994-2_COPY_613_752 x US-09-335-032-12207/rev ..
Align seg 1/1 to reverse of: US-09-335-032-12207 from: 1 to: 1531974

21 GlnHisSerLeuThrAsnSerProLeuAspSerCysLysGlu 37
|||||
59821 GGTGATTCAGAGCTACACCTCCCTGTCACACACATGATGCTATGC 59772
37 rSerValSerValThrSerProSerGlyVal.....SerSerThrS 52
|||||
59771 CAGTGTGAACACACATCTCTTACCAATATCATATGACATTCTAAAGCA 59722
52 ergLysGlyValSerSerThrSerAsnMetHisGlySerLeu...LeuGln 67
::|||
59721 ATACTGGTGCATCAGCAGCTGCACTGACATGACATGACATATCAACGCANA 59672
68 GluLysHisArgIleLeuHisLysLeuGlnAsnGlyAsnSerProAl 84
|||
59671 ACTACCACACCCGTTAGCCATGTCACCTGGTTCAGCAGCAGCTGCTTAA 59622
84 aglValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSerI 101
|||||
59621 ATCCACCGCTGTCGACAGGATCATTCAGTACTAAACTAACCCCTTCAG 59572
101 leThrSerCysGlyAspGlyAsnValAlaLysGlnGluInLeuSerPro 117
::|||
59571 TTACTCAT...GGACATACAACTGTT..... 59549
118 LysLysLysGluAsnAlaLeuLeuArgTyrLeuLeuAspArgAspAs 134
|||||
59548 .....AAGATTAACCTCTCAGCTACGAGGCTTATTTCGACACCTGATTC 59505
134 proSerAsp 137
::|||
59504 CAACCTCAGAT 59495

seq_name: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-032-12210
seq_documentation_block:
; Sequence 12210, Application US/09335032

```

```

; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; TITLE OF INVENTION: Transcriptome
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12210
; LENGTH: 1090936
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-12210

alignment_scores:
Quality: 87.50 Length: 166
Ratio: 0.941 Gaps: 10
Percent Similarity: 56.024 Percent Identity: 27.711

alignment_block:
US-09-041-994-2_COPY_613_752 x US-09-335-032-12210 ..
Align seg 1/1 to: US-09-335-032-12210 from: 1 to: 1090936

16 SerSerAspArgGlyHisSerSerLeuThrAsnSerPro.....Le 30
::|||
161167 AACAGTAATCTGTCAGAGCTCATCATCAATGAGCAATATCCGCGAGCT 161216
30 uAspSerSerCysLysGluSerSerValSerValThrSerProSerGly 47
|::|||
161217 CAGTGTATGCCCTCAAGAACAGCAATATA...TTAACTGCCCAAGCCGTTA 161263
47 aLerSerSerThrSerGlyGlyValSerSerThrSerAsn.....Met 61
::|||
161264 AAACAGACCATACACACAGCCGTAGATTAAGCTTCAATATTAATATAC 161313
62 HisGlySerLeuLeuGlnGluLys.....HisArgIleLe 73
|||||
161314 CATTAATAGCGCTGTTACATATGAAAGCTTTGAAGAAGATATGTGTGCT 161363
73 uHisLysLeuLeuGlnAsnGlyAsnSerProAla...GluValAlaLysI 89
::|||
161364 AGAGAAAGAAATCGGTTGAAGTTAATTCATTGGCAGACGAGTTCCTCA 161412
89 leThrAlaGlnAlaThrGlyLysAspThrSerSerIle.....Thr 102
::|||
161413 .....GCAGATTCAAATCCAAATCTATTATAGCACCCCACT 161448
103 SerCysGlyAspGlyAsnValAlaLysGlnGluInLeuSerProLys 119
|||
161449 TCACACCAAAATCAGACAGCTTATTGAATGAGCAATCTCTCCAAACA 161498
119 sLys.....GluAsnAsnAlaLeuLeuArg.... 127
::|||
161499 TCACACGATATTTTCAAAATCAAGAGAAATCCGAGTTTACGAGGCGCA 161548
128 .....TyrLeuLeu 130
::|||
161549 CATCATCTTCCAGTGGAGCTAGTATGGCTTAGCGCACCATCTTTGGTG 161598
131 AspArgAsp.....AspProSerAspAlaLeuSer 140
|||||
161599 GATAGACGCTGTCTATATACCTCGCTGAATCATCTTAATGCATTATCA 161646

seq_name: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-032-12210

```



```

seq_documentation_block:
; Sequence 12210, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Vogeliescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; TITLE OF INVENTION: Transcription
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; PRIOR FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 12210
; LENGTH: 1090936
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-12210

Alignment_scores:
Quality: 87.00 Length: 119
Ratio: 1.426 Gaps: 4
Percent Similarity: 51.261 Percent Identity: 29.412

Alignment_block:
US-09-041-994-2_COPY_613_752 x US-09-335-032-12210/rev ..
Align seg 1/1 to reverse of: US-09-335-032-12210 from: 1 to: 1090936

26 ThrAspSerProLeuAspSerSerCysIysGluSerValSerValThr 42
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
442260 ACCAATACACCTACCGTCGCTACGTCGACGTCGCTACTGCTCTAC 442211

42 rSerProGlyValSerSerThrSerThrSerGlyValSerSerThr 59
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
442210 TCTCTCTCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 442161

59 eAspMetHisGlySerLeuGluGlnIuGlyHisArgIleLeuHisIys 75
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
442160 CGCTCTACTTCTGCTCC..... 442144

76 LeuLeuGlnAsnGlyAsnSerProAlaGluValAlaIysIleThrAla 92
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
442143 .....TCTACTTCAATCCACACAGGAACACCCGCTACCACTCTGA 442103

92 nAlaThrGlyAspThrSerSerIleThrSerCysGlyAspGlyAsnV 109
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
442102 ACCTTCTCTCTGTACACCGCTGCTATTAATCTCT..... 442069

109 alValIysGlnGlnIleuSerProIysIysGluAsnAlaLeu 125
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
442068 .....TCTCCAAG.....GCT 442057

126 LeuAlaGlyTyrLeuLeuAspArgAspAspProSer.....AspAl 138
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
442056 AATTCATATTTCACCATTAATGAATGAGGCTCTGTAACTGTCAAGCGC 442007

138 alLeuSer 140
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
442006 TGTCTCT 442000

seq_name: /cgn2_e/ptodata/2/pna/PCT_NEW_COMB.seq:PCT-US01-01339-9282
seq_documentation_block:
; Sequence 9282, Application PC/TUS0101339
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences Inc., et al.

```

```

; TITLE OR INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006PCr
; CURRENT APPLICATION NUMBER: PCT/US01/01339
; CURRENT FILING DATE: 2001-03-17
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9282
; LENGTH: 5402
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (4456)
; OTHER INFORMATION: n equals a,t,g, or c
PCT-US01-01339-9282

alignment_scores:
      Quality:      85.50      Length:      158
      Ratio:        0.983      Gaps:         8
      Percent Similarity: 55.063      Percent Identity: 27.848

alignment_block:
US-09-041-994-2_COPY_613_752 x PCT-US01-01339-9282/rev ..

Align seg 1/1 to reverse of: PCT-US01-01339-9282 from: 1 to: 5402

      9 LeuLeuGlnLeuLeuThrcYsserSeraspasparglyHisserSerle 25
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2199 CTAATGCCAGCGCTGAGTGCAGTAGTACGATGTTGGCTCAGTCGACATCT 2150
25 u.....ThraSerPro.....LeuAspSerSerCysl 35
      |
      |GCGCTCCAGCGCTCAGCAGTGCATGCTCCAGATAGCTGAGACGAC 2140
35 ysgLserSerValserValThrserProserglyValserSerSerThr 51
      |||||:|||||:||||| ||||| |||||:|||||:
2099 ACCGCAAGCAAGTACTGCTGGAGCTCAGCGCTCCAGTAGCTGGGACCTCA 2050
52 SerglyValserSerThr.....SerasmethisglSerleuLeu 66
      ::::: |||||:|||||: |||||: |||||: |||||: |||||:
2049 GCCTCCAGATGCTGGGAGACTACAGCCACACACACCATGGCCAGCTAAT 2000
66 u.....GlnGluLysHisArgIleL 73
      :
      :TTTTGATTTCTTGTAGAGATGGGGTTTGGCATGTGGCCAC..GCTG 1993
73 euHisLysLeuLeuGlnasnGlyasnSerProalagluValAlaLysIle 89
      |||||:|||||: |||||: |||||: |||||: |||||: |||||:
1952 GTCATGTAAGCTACGAGCTCAAGCAATGTCACACACTGCTGGTCCCAAGTG 1903
90 Thr.....AlaGlnAlaThrGlyLysAspHisSerSerI 101
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
1902 TTGGGATTACAGCATGTGAGCCACTGCGGCTGGTAGAAGACTGATTTT 1853
101 eThrSerCysglAspGlyasn.....Val.ValLys 111
      :|||:|||||: |||||: |||||: |||||: |||||: |||||:
1852 AACAAAGCTTCCACAGATGGTTTCAAATGCAACACTTCAGTTTGTAACATG 1803
112 GlnGluGlnLeuSerProLysLysGlnasnAlaLeuLeuArgTy 128
      ::::: |||||: |||||: |||||: |||||: |||||: |||||:
1802 TCCTGTAGAGGCTCTCTATGATGATGCTTCCAGCAGCAATATGCTCTCATTTTA 1753
128 IleuLeuAspArgAspAspPro 135
      |||||:|||||: |||||: |||||: |||||: |||||: |||||:
1752 TCTCTGTCCAAAGATTTCCCA 1731

seq_name: /cgn2_6/prodata/2/pna/US09_NEW_COMB.seq:US-09-335-032-12214

seq_documentation_block:
; Sequence 12214, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor

```

```
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12214
; LENGTH: 666448
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-12214
```

```
alignment_scores:
  Quality: 85.50      Length: 118
  Ratio: 1.082      Gaps: 2
  Percent Similarity: 66.949      Percent Identity: 23.729
```

## alignment\_block:

US-09-041-994-2\_COPY\_613\_752 x US-09-335-032-12214/rev ..

Align seg 1/1 to reverse of: US-09-335-032-12214 from: 1 to: 666448

```
18 AspaAspArgGlyHisSerSerLeuThrAsnSerProLeuAspSerSerCy 34
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
613459 GAAGAGAAATCTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 613410
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
34 slyGluSerSerValSerValThrSerProSerGlyValSerSerSer 51
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
613409 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 613360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
51 hrSerGlyGlyValSerSerThrSerAsnMetHisGlySerLeuGln 67
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
613359 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGT 613310
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
68 GluLysHisArgIleLeuHisLysLeuGlnAsnGlyAsnSerProAl 84
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
613309 GACTCTGAATCATGCTCATCATCTTCTCTCTCTCTCTCTCTCTCTCAG 613260
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
84 a.....GluValAlaLysIleThrAlaGlnAlaIleThrGlyL 96
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
613259 TTCACAGTCACTGGAATCTTCTAGTGAATCCGACTCTAGCTCCAGTGGAT 613210
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
96 yAspThrSerSerIleThrSerCysGlyAspGlyAsnValIleLysGln 112
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
613209 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCGAA 613160
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
113 .GluGlnLeuSerProLysLysLysGluAsnAsnAlaLeuLeuArgTyrL 129
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
613159 TCAGAAATGATAACCAAGAAAGAGCAAGAAATGACACAAACAGAAATGTC 613110
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
129 eu 129
   ||
613109 TA 613108
```

seq\_name: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-335-032-12212

## seq\_documentation\_block:

```
; Sequence 12212, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the yeast
; FILE REFERENCE: 01107.78572
```

```
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12212
; LENGTH: 439885
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-12212
```

```
alignment_scores:
  Quality: 85.00      Length: 105
  Ratio: 1.232      Gaps: 3
  Percent Similarity: 65.714      Percent Identity: 28.571
```

## alignment\_block:

US-09-041-994-2\_COPY\_613\_752 x US-09-335-032-12212/rev ..

Align seg 1/1 to reverse of: US-09-335-032-12212 from: 1 to: 439885

```
23 SerSerLeuThrAsnSerProLeuAspSerSerCysLysGluSerSerVa 39
   |||||: : : : : : : : : : : : : : : : : : : : : : : :
25662 AGCTCTCAGTCAGTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 25613
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
39 l.....SerValThrSerProSerGlyValSerSerSerThrSerCylg 54
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
25612 CAGTCATCAAGCTCCCTCCAGCTTCTGATGTCTCAAGCTCTGTACGTCAAT 25563
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
54 lylValSerSerThrSerAsnMetHisGlySerLeuGlnGlyLysHis 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
25562 CAGCTTCTCTACTTCTGATGTCTCAAGCTCTGTCAGTCATCTTCTTCT 25513
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
71 Arg.....IleLeuHisLysLeuGlnAsnGlyAsnSerProAl 84
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
25512 TCAGCTCTGTGATGATCAAGCTCTGTGATGATCAATCTTCTCTCAGCTTTC 25463
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
84 aglValAlaLysIleThrAlaGlnAla..... 93
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
25462 TGATGTCTCAAGCTCTGTGATGATCAATCAAGCTCTCTCTCTCTCTCTCT 25413
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
94 .....ThrGlyLysAspThrSerSerIleThrSerCys 104
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
25412 CAAGCTCTGTCAAGTCATCAAGCTTCTCTCTCTCTCTCTCTCTCTCTCT 25363
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
105 glyAspGlyAsnVal 109
   ||||: : : : :
25362 GGCAGCCAAATCAGTC 25348
```

seq\_name: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-335-032-12218

## seq\_documentation\_block:

```
; Sequence 12218, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12218
```

LENGTH: 1091282  
TYPE: DNA  
ORGANISM: Saccharomyces cerevisiae  
US-09-335-032-12218

alignment\_scores:  
Quality: 83.00 Length: 102  
Ratio: 1.297 Gaps: 4  
Percent Similarity: 62.745 Percent Identity: 34.314

## Alignment block:

US-09-041-994-2\_COPY\_613\_752 x US-09-335-032-12218 ..

Align seg 1/1 to: US-09-335-032-12218 from: 1 to: 1091282

```
16 SerSerAspArgGlyHisSerSerLeuThrAsnSerProLeuAspSe 32
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
427909 TCGTGTGTCGTCGTCGTCATCATCGTCATCATCGTCATCATC 427958
32 rSerCysLysGlnSerSerValThrSerProSerGlyValSerS 49
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
427959 ATCATTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCATCAT 428008
49 eSerThrSerGlyValSerSerThrSerAsnMetHisGlySer... 64
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
428009 CGTCATCATCATCATCGTCATCGTCGTCGTCATCATCATCTTGA 428058
65 .....LeuLeuGlnGlnLysHisArgIleLeuHisL 75
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
428059 CCACGAGATGCTCTCTTATATCATCTCTCATGCTATCTTCATC 428108
75 ySLeuLeu.....GlnAsnGlnSerProAlaGlnValAla 87
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
428109 CTCCTTCCTCTCTTCCCGAGAAACATCGGATTAACACCTGATTGA 428158
88 LysIleThrAlaGlnAlaThrGlyLysAspThrSerSer...IleThrSe 103
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
428159 CGACTATCTTCCCAACATCATCTTGCATCCATCATCCCTTAACCG 428208
103 rCys 104
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
428209 ATGTT 428212
```

seq\_name: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-335-032-12212

seq\_documentation\_block:  
; Sequence 12212, Application US/09335032  
; GENERAL INFORMATION:  
; APPLICANT: Velculescu, Victor  
; APPLICANT: Vogelstein, Bert  
; APPLICANT: Kinzler, Kenneth  
; TITLE OF INVENTION: Characterization of the Yeast  
; FILE REFERENCE: 01107.78572  
; CURRENT APPLICATION NUMBER: US/09/335,032  
; CURRENT FILING DATE: 1999-06-16  
; PRIOR APPLICATION NUMBER: US 60/035,917  
; PRIOR FILING DATE: 1997-01-23  
; PRIOR APPLICATION NUMBER: US 09/012,031  
; PRIOR FILING DATE: 1998-01-22  
; NUMBER OF SEQ ID NOS: 12219  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12212  
; LENGTH: 439885  
; TYPE: DNA  
; ORGANISM: Saccharomyces cerevisiae  
US-09-335-032-12212

alignment\_scores:  
Quality: 82.00 Length: 111  
Ratio: 1.079 Gaps: 0

Percent Similarity: 68.468 Percent Identity: 24.324

## Alignment block:

US-09-041-994-2\_COPY\_613\_752 x US-09-335-032-12212 ..

Align seg 1/1 to: US-09-335-032-12212 from: 1 to: 439885

```
14 ThrCysSerSerAspArgGlyHisSerSerLeuThrAsnSerProLe 30
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
334123 ACTGCTTCCTCAGAGCGCTTCTGCTTGAACCTTCAGCGCTTC 334172
30 uAspSerCysLysGlnSerSerValThrSerProSerGlyVal 47
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
334173 TTCTTTCGCTTCTGATTCCTTTCAGCGCTTCCTCTCTCTGAT 334222
47 aSerSerSerThrSerGlyValSerSerThrSerAsnMetHisGly 63
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
334223 CCTCTCAGCGCTTCTCTCTCTCTCAGATCTCTCTGCGCTG 334272
64 SerLeuLeuGlnGlnLysHisArgIleLeuHisLysLeuLeuGln 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
334273 TCTTCTGCTTCTGAGAGCTGTAAGCTTCTGAGCTGCCAAGTCTTG 334322
80 yAsnSerProAlaGlnValAlaLysIleThrAlaGlnAlaThrGlyLys 97
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
334323 CTCCTTCGCTCTTCATCATGCTGCTTCATCTGCTTCTTCCAGCGCT 334372
97 sPThrSerSerIleThrSerCysGlnAspGlnValValLysGlnGlu 113
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
334373 CTGACGCTTCTCTCTGCAAAAGCGCTCTCTCTGCGAGAAAATCTACT 334422
114 GlnLeuSerProLysLysGlnAsnAsnAla 124
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
334423 AATGAGCTCTCTCTGCTGCTACCTCCAGAACGCT 334455
```

seq\_name: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-335-032-12208

seq\_documentation\_block:  
; Sequence 12208, Application US/09335032  
; GENERAL INFORMATION:  
; APPLICANT: Velculescu, Victor  
; APPLICANT: Vogelstein, Bert  
; APPLICANT: Kinzler, Kenneth  
; TITLE OF INVENTION: Characterization of the Yeast  
; FILE REFERENCE: 01107.78572  
; CURRENT APPLICATION NUMBER: US/09/335,032  
; CURRENT FILING DATE: 1999-06-16  
; PRIOR APPLICATION NUMBER: US 60/035,917  
; PRIOR FILING DATE: 1997-01-23  
; PRIOR APPLICATION NUMBER: US 09/012,031  
; PRIOR FILING DATE: 1998-01-22  
; NUMBER OF SEQ ID NOS: 12219  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12208  
; LENGTH: 574860  
; TYPE: DNA  
; ORGANISM: Saccharomyces cerevisiae  
US-09-335-032-12208

alignment\_scores:  
Quality: 80.00 Length: 95  
Ratio: 1.356 Gaps: 2  
Percent Similarity: 62.105 Percent Identity: 32.632

## Alignment block:

US-09-041-994-2\_COPY\_613\_752 x US-09-335-032-12208 ..

Align seg 1/1 to: US-09-335-032-12208 from: 1 to: 574860

```
16 SerSerAspArgGlyHisSerSerLeuThrAsnSerProLeuAspSe 32
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

```

79130 TCTTCTGACACGACCTCTTCTCCAAAGTCATCGATCAGACGAG 79179
32 rSerCysLysGluSerSer.....ValSerValThrSerProSerGlyV 47
||||| ||||| ||||| ||||| |||||
79180 TTCACTAGTAAGAGTAGTAGTAGTAGTTCACCGCTTCATCATCATCAT 79229
47 a1SerSerThrSerGlyValSerSerThrSerAsnMetHisGly 63
||||| ||||| ||||| ||||| |||||
79230 CATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 79279
64 SerLeuLeuGlnGlnLysHisArgLeuHisLysLeuLeuGlnAsnG1 80
::: ||||| ||||| ||||| ||||| |||||
79280 GAT.....AAAGTCGATATCTCCGT 79299
80 yAsnSerProAlaGluValAlaLysIleThrAlaGlnAlaThrGlyLysA 97
::: ||||| ||||| ||||| ||||| |||||
79300 ATCCCTTCCGCTCACTCCCAACATCAACACATCTCTCTGTAACGGGA 79349
97 sPThrSerSerIleThrSerCysGlyAspGlyAsn 108
::: ||||| ||||| ||||| ||||| |||||
79350 GCGCCTCCAGCAGTAGTACAGCTCAATGACGCGTAAT 79384

seq_name: /cgn2_6/p/odata/2/pna/us08_NEW_COMB.seq:us-08-956-171C-480

seq_documentation_block:
: Sequence 480, Application US/08956171C
: GENERAL INFORMATION:
: APPLICANT: Charles Kunsch
:      Gil H. Choi
:      Patrick S. Dillon
:      Craig A. Rosen
:      Steven C. Barash
:      Michael R. Fannon
: TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
: NUMBER OF SEQUENCES: 5255
: CORRESPONDENCE ADDRESS:
:      ADDRESSEE: Human Genome Sciences, Inc.
:      STREET: 9410 Key West Avenue
:      CITY: Rockville
:      STATE: Maryland
:      COUNTRY: USA
:      ZIP: 20850
: COMPUTER READABLE FORM:
:      MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
:      COMPUTER: HP Vectra 486/33
:      OPERATING SYSTEM: MSDOS version 6.2
:      SOFTWARE: ASCII Text
: CURRENT APPLICATION DATA:
:      APPLICATION NUMBER: US/08/956,171C
:      FILING DATE: 20-Oct-1997
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
:      APPLICATION NUMBER: 60/009,861
:      FILING DATE: January 5, 1996
:      APPLICATION NUMBER: 08/781,986
:      FILING DATE: January 3, 1997
: ATTORNEY/AGENT INFORMATION:
:      NAME: Hoover, Kenley K.
:      REGISTRATION NUMBER: 40,302
:      REFERENCE/DOCKET NUMBER: PB248P1
: TELECOMMUNICATION INFORMATION:
:      TELEPHONE: (301) 610-5790
:      TELEFAX: (301) 309-8439
: INFORMATION FOR SEQ ID NO: 480:
:      SEQUENCE CHARACTERISTICS:
:          LENGTH: 6309 base pairs
:          TYPE: nucleic acid
:          STRANDEDNESS: double
:          TOPOLOGY: linear
:      SEQUENCE DESCRIPTION: SEQ ID NO: 480:
US-08-956-171C-480

```

```

alignment_scores:
  Quality: 78.50      Length: 164
  Ratio: 0.946      Gaps: 6
  Percent Similarity: 50.610      Percent Identity: 22.561

alignment_block:
US-09-041-994-2_COPY_613_752 x US-08-956-171C-480 ..
Align seg 1/1 to: US-08-956-171C-480 from: 1 to: 6309

6 HisLysLysLeuLeuGlnLeuLeuThrCysSerSerAspAspArgGlyH1 22
||||| ||||| ||||| ||||| |||||
82 CATAGTAGATTAATAGAAAGACACCAAAATTAACGATGACTTAGGTGG 131
22 sSerSerLeuThrAsnSerProLeuAspSerSerCysLysGluSerServ 39
::: ||||| ||||| ||||| ||||| |||||
132 CGGTTCATTAATGATGATTAATGCAATTAATGCAACA...CAAGTGGTGATA 178
39 a1SerValThrSerProSerGlyValSerSerThrSerGlyLysVal 55
::: ||||| ||||| ||||| ||||| |||||
179 TTTCAGCTTAATGATGACCAACAAAGCTTAATTCATTAATGACATGCAAAATC 228
55 .....LeuLeuGlnLysHisArgLeuLeu 73
329 AGAAAGTTGCTGCTGATGCTTACGCTTACCTTACCTTACCTTACCTTAC 378
74 HisLysLeuLeuGlnLysAsnSerProAlaGlnAlaThrGlyLysA 90
::: ||||| ||||| ||||| ||||| |||||
379 GAAATCTTTGCACAATTCGCT.....TCAGACTTGTGATGAATTTAC 419
90 rAlaGlnAlaThrGlyLysAspThrSerSerIleThrSerCysGlyAspG 107
||||| ||||| ||||| ||||| |||||
420 TCCAAGTAATATAGAACGCGTAAACGCTT..... 453
107 LysValValLysGlnGlnLeuSerProLysLysLysLysAsn 123
::: ||||| ||||| ||||| ||||| |||||
454 ..GAAGCTTAAACAAAGATCAAAACCAATTCACACTCAGCTCAACACCA 501
124 AlaLeuLeuArgTyrLeuLeuAspArg.....AspAsp 134
::: ||||| ||||| ||||| ||||| |||||
502 GTGTGATTAATTAATGATTAACAAAGATATTAGATGAT 543

seq_name: /cgn2_6/p/odata/2/pna/us08_NEW_COMB.seq:us-08-956-171C-220

seq_documentation_block:
: Sequence 220, Application US/08956171C
: GENERAL INFORMATION:
: APPLICANT: Charles Kunsch
:      Gil H. Choi
:      Patrick S. Dillon
:      Craig A. Rosen
:      Steven C. Barash
:      Michael R. Fannon
: TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
: NUMBER OF SEQUENCES: 5255
: CORRESPONDENCE ADDRESS:
:      ADDRESSEE: Human Genome Sciences, Inc.
:      STREET: 9410 Key West Avenue
:      CITY: Rockville
:      STATE: Maryland
:      COUNTRY: USA
:      ZIP: 20850
: COMPUTER READABLE FORM:
:      MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
:      COMPUTER: HP Vectra 486/33

```



```

      |||
95162 CTTGAAGCCCGCAGCTGACGCTCCGATCCCGCCGAGGTTTA 95207
      |||
seq_name: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:us-09-335-032-12216
seq_documentation_block:
: Sequence 12216, Application US/09335032
: GENERAL INFORMATION:
: APPLICANT: Velculescu, Victor
: APPLICANT: Vogelstein, Bert
: APPLICANT: Kinzler, Kenneth
: TITLE OF INVENTION: Characterization of the Yeast
: TITLE OF INVENTION: Transcriptome
: FILE REFERENCE: 01107.78572
: CURRENT APPLICATION NUMBER: US/09/335.032
: CURRENT FILING DATE: 1999-06-16
: PRIOR APPLICATION NUMBER: US 60/035,917
: PRIOR FILING DATE: 1997-01-23
: PRIOR APPLICATION NUMBER: US 09/012,031
: PRIOR FILING DATE: 1998-01-22
: NUMBER OF SEQ ID NOS: 12219
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 12216
: LENGTH: 924430
: TYPE: DNA
: ORGANISM: Saccharomyces cerevisiae
US-09-335-032-12216

alignment_scores:
      Quality: 78.00      Length: 100
      Ratio: 1.418      Gaps: 3
      Percent Similarity: 55.000      Percent Identity: 30.000

alignment_block:
US-09-041-994-2_copy_613_752 x US-09-335-032-12216/rev ..
Align seg 1/1 to reverse of: US-09-335-032-12216 from: 1 to: 924430

      23 SerSerLeuThrAsnSerProLeuAspSerSerCysLysGluSerSerGly 39
      ||||| :|||:||||| :|||:||||| :|||:|||||
879914 TCTTCAGATTCATCTCACCCGCTCATCTCAATTCATCTTCAGC 879865
      39 lSerValThrSerProSerGlyValSerSerSerThrSerGlyValS 56
      :||| :|||:||||| :|||:||||| :|||:|||||
879864 TTCATCTCCGCTTCATCTAGTATCTCAGCGCTCCGCTCAAGGCTATTA 879815
      56 er...SerThrSerAsnMetHisGlySerLeuGluGlnIuLysHisArg 71
      :||| :||| :|||:||||| :|||:|||||
879814 CTTCAGCTCTTCATCATGATGAGGCTCCGT..... 879783
      72 lLeuHisLysLeuLeuGlnAsnGlyAsnSerProAlaGluValAla 88
      :||| :|||:||||| :|||:||||| :|||:|||||
879782 .....AAATCTACTGCTCAAGTCCGCTC 879760
      88 glIleThrAlaGlnAlaThrGlyLysAspThrSerSerIle.....Ths 103
      :||| :|||:||||| :|||:||||| :|||:|||||
879755 AGATTAGAACAGCTGATGCTGTTTGAACAACATCAGATTATATGGCGT 879710
      103 erCysGlyAspGlyAsnValLysGlnGluGlnLeuSerProLysLys 119
      :|||:||||| :|||:||||| :|||:||||| :|||:|||||
879709 ACTGTATGTCAGGTGATGATGTTCTTGCACAGCTTAACCTCAAGCCAGAAA 879660

seq_name: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:us-09-335-032-12219
seq_documentation_block:
: Sequence 12219, Application US/09335032
: GENERAL INFORMATION:
: APPLICANT: Velculescu, Victor
: APPLICANT: Vogelstein, Bert
: APPLICANT: Kinzler, Kenneth
: TITLE OF INVENTION: Characterization of the Yeast
: TITLE OF INVENTION: Transcriptome
```

```

: FILE REFERENCE: 01107.78572
: CURRENT APPLICATION NUMBER: US/09/335.032
: CURRENT FILING DATE: 1999-06-16
: PRIOR APPLICATION NUMBER: US 60/035,917
: PRIOR FILING DATE: 1997-01-23
: PRIOR APPLICATION NUMBER: US 09/012,031
: PRIOR FILING DATE: 1998-01-22
: NUMBER OF SEQ ID NOS: 12219
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 12219
: LENGTH: 948061
: TYPE: DNA
: ORGANISM: Saccharomyces cerevisiae
US-09-335-032-12219

alignment_scores:
      Quality: 79.50      Length: 129
      Ratio: 1.187      Gaps: 7
      Percent Similarity: 51.938      Percent Identity: 29.457

alignment_block:
US-09-041-994-2_copy_613_752 x US-09-335-032-12219/rev ..
Align seg 1/1 to reverse of: US-09-335-032-12219 from: 1 to: 948061

      29 ProLeuAspSerSerCysLysGluSerValSerValThrSerProSe 45
      ||| :||| :|||:||||| :|||:|||||
295710 CCTGATGACATGATTAAGCGAGTCTTCATTTCCGTG..... 295672
      45 rGlyValSerSerThrSerThrSerGlyValSerSerThrSerAsnMet 62
      :|||:||||| :|||:||||| :|||:||||| :|||:|||||
295671 .GCTATTTCAGTAGACACCAATCGGCCCAATGCCAATTCACACATTA 295623
      62 lsgLysSerLeuGluGlnGluLysHisArgIleuHisLysLeuGln 78
      :||| :||| :|||:||||| :|||:|||||
295622 AGGACCGCTGTACTAAGAACTTAAGAAATATATCTCGCCCTGGA 295573
      79 Asn..... 79
      |||
295572 AATCAAAAGTTGGCTTTAAGGAGCGCTGTAAGCATCAGCAAGAACTCA 295523
      80 .....GlyAsnSerProAla...GluValAla.....LysIleThr 91
      :||| :|||:||||| :|||:||||| :|||:|||||
295522 TTTTGTGCTTAATAGAAATAGCCCTTGAAGTGAATTTTAAGACTGCTT 295473
      91 lacIlnAlaThrGlyLysAspThrSerSerIleThrSerCysGlyAspGly 107
      :|||:||||| :|||:||||| :|||:||||| :|||:|||||
295472 CGGATGACAGCGGTTAAGCTTACTGAA.....AAGTCATCAGAAAGC 295432
      108 AsnValValLysGln...GluGlnLeuSerPro..... 117
      :|||:||||| :|||:||||| :|||:||||| :|||:|||||
295431 ACCATTACGAAGCAAGAAAGAGAGTGTCACCTCCCAATATGTTGAAGCTGA 295382
      118 .LysLysLysGluAsnAsnAlaLeuLeuArgGlyLeu 129
      :|||:||||| :|||:||||| :|||:||||| :|||:|||||
295381 TCGGAGCTTGAGAAATTAATTAACCTTCAGAAAAACCTC 295345
```



This Page Blank (uspto)